DOCUMENT RESUME

ED 354 684 EC 301 867

AUTHOR Hinton, Daniel E., Sr.
TITLE Examining Advanced Technologies for Benefits to Persons with Sensory Impairments. Final Report.
INSTITUTION Science Applications International Corp., Arlington, VA.
SPONS AGENCY Special Education Programs (ED/OSERS), Washington, DC.
REPORT NO SAIC-92/1059
PUB DATE Mar 92
CONTRACT HS90047001
NOTE 397p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC16 Plus Postage.
DESCRIPTORS *Accessibility (for Disabled); Assistive Devices (for Disabled); Blindness; Braille; *Communication Aids (for Disabled); Computer Oriented Programs; Computers; Deafness; *Hearing Impairments; Input Output Devices; Interactive Video; Sensory Aids; *Technological Advancement; *Telecommunications; *Visual Impairments

ABSTRACT

This final report describes activities and products of an 18-month study on improving access of persons with sensory impairments to media, telecommunications, electronic correspondence, and other communications devices by means of technological advancements. Ten scenarios were developed which describe potential applications of: (1) Braille devices (a major technology shift is required to design a full-page Braille display); (2) input/output devices (alternative display or translator devices to computers are needed); (3) visible light spectrum manipulation (computer access, night vision, and image enhancement technology can improve access to printed media); (4) flat panel displays (key technologies include hand-held or flatbed scanners with optical character reader software); (5) descriptive video (costs seem to be the barrier to implementation); (6) adaptive modems (combined access to telephone devices for the deaf and data transmission over telephone lines is needed); (7) telecommunications systems (development of voice recognition systems, call progress information, and access to automatic message answering systems is needed); (8) voice recognition systems (voice recognition systems that are speaker independent are required); (9) video teleconferencing/data compression (true video phones--but not picture phones--are attractive to deaf persons); and (10) portable power systems (selection of appropriate power sources is urged). Appendices, which constitute the bulk of this report, include a conceptual framework document, the information collecting plan, a 10-year development plan, and the full text of all 10 scenario papers. (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

ED354684

FINAL REPORT
for
"EXAMINING ADVANCED TECHNOLOGIES FOR BENEFITS TO PERSONS WITH SENSORY IMPAIRMENTS"

MARCH 1992

Report Number: SAIC-92/1059

Department of Education Contract # HS90047001

Prepared By

DANIEL E. HINTON, SR.
SCIENCE APPLICATIONS INTERNATIONAL CORPORATION
3701 N. Fairfax Drive, Suite 1001
Arlington, VA 22203
(703) 351-7755

for

DEPARTMENT OF EDUCATION
OFFICE OF SPECIAL EDUCATION PROGRAMS
300 C Street, S.W.
Washington, DC 20202

EC 301867

BEST COPY AVAILABLE

TABLE OF CONTENTS

Section Page

Acknowledgement vi
Glossary vii
Executive Summary viii

1.0 INTRODUCTION 1
1.1 Report Structure 1
1.2 Background 2

2.0 PURPOSE AND OBJECTIVE 3
2.1 Purpose 3
2.2 Objectives 3
2.3 Program Tasks 3

3.0 TECHNICAL APPROACH 5
3.1 Program Management and Control Tasks 5
3.2 Expert Advice and Oversight Tasks 6
3.3 Study Execution Tasks 6

4.0 RESULTS AND FINDINGS 12
4.1 Technologies for Visual Impairments 12
4.1.1 Braille Devices and Techniques to Allow Media Access 12
4.1.2 Input/Output Devices for Computer and Electronic Books Access 15
4.1.3 Visible Light Spectrum Manipulation to Allow Media for Persons with Selective Vision 19
4.1.4 Character Readers for Dynamic LED and LCD Display Access 22
4.1.5 Descriptive Video for TV Access 25
4.2 Technologies for Hearing Impairments 30
4.2.1 Adaptive Modems and TDD Access 30
4.2.2 Telecommunications System Access 35
4.2.3 Voice Recognition Systems for Personal and Media Access 38
4.2.4 Video Teleconferencing/Data Compression for Persons with Hearing Impairments 42
4.3 Technologies for Visual and/or Hearing Impairments 46

5.0 CONCLUSIONS 49
5.1 Technologies for Visual Impairments 49

TABLE OF CONTENTS (Continued)

Section Page

5.1.1 Braille Devices and Techniques to Allow Media Access 49
5.1.2 Input/Output Devices for Computer and Electronic Book Access 49
5.1.3 Visible Light Spectrum Manipulation to Allow Media Access for Persons with Selective Vision 49
5.1.4 Flat Panel Terminal Displays Used with Page Scanners 50
5.1.5 Descriptive Video for Television Access 50
5.2 Technologies for Hearing Impairment 51
5.2.1 Adaptive Modems and TDD Access 51
5.2.2 Telecommunications System Access 51
5.2.3 Voice Recognition Systems for Personal and Media Access 52
5.2.4 Video Teleconferencing/Data Compression for Persons with Hearing Impairments 52
5.3 Technology for Visual and/or Hearing Impairments 56

6.0 RECOMMENDATIONS 57
6.1 Technologies for Visual Impairment 57
6.1.1 Braille Devices and Techniques to Allow Media Access 57
6.1.2 Input/Output Devices for Computer and Electronic Book Access 57
6.1.3 Visible Light Spectrum Manipulation to Allow Media Access for Persons with Selective Vision 58
6.1.4 Flat Panel Terminal Displays Used with Page Scanners 58
6.1.5 Descriptive Video for TV Access 59
6.2 Technologies for Hearing Impairment 59
6.2.1 Adaptive Modems and TDD Access 59
6.2.2 Telecommunications System Access 61
6.2.3 Voice Recognition Systems for Personal and Media Access 62
6.2.4 Video Teleconferencing/Data Compression for Persons with Hearing Impairments 63
6.3 Technology for Visual and/or Hearing Impairments 64

TABLE OF CONTENTS (Continued)

Appendices Page

A CONCEPTUAL FRAMEWORK DOCUMENT A-1

B INFORMATION COLLECTION PLAN B-1

C TEN-YEAR DEVELOPMENT PLAN C-1

D SCENARIOS D-1

LIST OF ILLUSTRATIONS

Table Page

3.2-1 Distinguished Panel of Experts 7
3.3-1 List of Scenarios 8
3.3-2 List of Deliverables 11

4.1.3-1 Alternate Display Systems Usable with All Software 20
4.1.5-1 Potential DVS Users by Level of Visual Impairment 26
4.1.5-2 Technologies Capable of Broadcasting Described Video 27
4.3-1 Commercial Batteries 47

Figure Page

3.3-1 Program Conceptual Framework 10

4.2.4-1 Taxonomy of Compression Algorithms 45

ACKNOWLEDGEMENTS

The authors of this report acknowledge the support provided by Mr. Ernie Hairston of the U.S. Department of Education, Office of Special Education Programs, throughout the program. Without his oversight and assistance, this important work would not have been possible.

The Panel of Experts proved their value throughout the 19 months of the program. Their dedication to meeting the media access needs of persons with sensory impairments was an inspiration to the SAIC staff and was invaluable in our research, information collection, and scenario development efforts.

The Conference Center staff is to be commended for planning and execution of the Panel of Experts meetings. Dr. Carl Jensema also provided exceptional input into each scenario in the area of hearing impairments. The team of a small business doing specialized conference administration and a large technical corporation proved to be the optimum mix of talent for this major advanced technology research program.

Finally, the efforts of the SAIC technical and administrative staff were exemplary. The engineers volunteered extra time on each scenario. Mr. Charles Connolly (principal writer), Mr. Paolo Basso-Luca, Mr. Lewis On, Mr. Rainer Kohler, and Mr. Daniel Morrison all made significant contributions to this work. Mrs. Nancy Davis, as the administrative assistant, helped write two scenarios as well as provided the administrative and technical expertise in drafting, editing and finalizing the document. Mr. John Park, a senior SAIC staff scientist, voluntarily edited the document on his own time.

The report that follows is a reflection of the dedication, skill and teamwork that is possible when the Government, industry and the academic community work as a cooperative team.

GLOSSARY

ADA	Americans for Disability Act
AFB	American Foundation for the Blind
ARS	Adverse-Environment Recognition of Speech
ASICS	Application-specific integrated circuits
ASL	American Sign Language
ATIS	Air Travel Information System
ATV	Advanced television
BAM	Bit assignment matrices
B-ISDN	Broadband Integrated Services Digital Network
bps	bit-per-second
BTSC	Broadcast Television Systems Committee
CCD	Charge-coupled device
CD	Continuous density
CNN	Cable News Network
CSR	Continuous speech recognition
DARPA	Defense Advanced Research Project Agency
DCT	Discrete cosine transform
DOD	Department of Defense
DPCM	Differential pulse code modulation
DSP	Digital signal processors
DTMF	Dual tone multi-frequency
DTW	Dynamic Time Warp
DV	Descriptive video
DVS	Descriptive Video Services
EIA	Electronic Industries Association
EIF	Electronic Industries Foundation
ER	Electrorheological
ESPRIT	European Strategic Programme for Research and Development in Information Technology
FCC	Federal Communications Commission
FFT	Fast Fourier Transform
HBO	Home Box Office
HDTV	High definition television
HMM	Hidden Markov Model
HUD	Head-up displays
ICP	Information Collection Plan
IDPCM	Interpolative differential pulse code modulation
I/O	Input/output
ISDN	Integrated Services Digital Network
IT	Information technology
LAN	Local area network
LCD	Liquid crystal displays
LED	Light emitting diode
LMS	Least mean squares
LPC	Linear predictive coding
MCU	Microcontroller unit
MIT	Massachusetts Institute of Technology
MTS	Multichannel Television Sound
NABTS	North American Basic Teletext Specification
NASA	National Aeronautics and Space Administration
N-ISDN	Narrowband Integrated Services Digital Network
NFB	National Federation of the Blind
NIDRR	National Institute for Disability and Rehabilitation Research
NIST	National Institute of Standards and Technology
NITF	National Image Transmission Format
NLS	National Library Service for the Blind and Physically Handicapped
NTSC	National Television Systems Committee
OAG	Official Airline Guide
OCR	Optical character recognition
OCV	Open circuit voltage
OSEP	U.S. Department of Education Office of Special Education Programs
PBS	Public Broadcasting System
PBX	Public Switch Exchange
PC	Personal computer
PCM	Pulse code modulation
Pro	Professional
PVDF	Polyvinylidene difluoride
PZT	Lead zirconate titanate
RFP	Request for Proposal
RM	Resource management
ROM	Read-only memory
RSA	Recognizer Sensitivity Analysis
SAIC	Science Applications International Corporation
SAP	Second audio program
SBIR	Small Business Innovative Research Programs
SCA	Subcarriers (radio)
SLI	Starting, lighting, ignition
SMPTE	Society of Motion Picture and Television Engineers
SNN	Segmental neural network
SRI	Standford Research Institute
TDD	Telephone device for the deaf
VBI	Vertical blanking interval
VDC	Volts direct current
VQ	Vector quantization
wpm	Words per minute
WST	World System Teletext

EXECUTIVE SUMMARY

Introduction

This final report covers the study performed by SAIC on "Examining Advanced Technologies for Benefits to Persons with Sensory Impairments." The study was conducted between October 1, 1990 and March 3, 1992 for the U.S. Department of Education Office of Special Education Programs (OSEP) under contract HS90047001.

Purpose and Objectives

The study contract was instituted by OSEP as a first step in a comprehensive media access advanced technology program to benefit persons with sensory impairments. The objectives of the technical study were to identify advanced or emerging technologies that might have applications that would facilitate the access of individuals with sensory impairments to media, telecommunications devices, electronic correspondence, and innovative uses of current communications devices. Also, the study identified activities that would be required to adopt and develop these technologies for the benefit of the sensory impaired.

Technical Results and Findings

A 15-member panel of experts on technology and sensory impairments provided guidance to the technical efforts by defining the areas of media and telecommunications access required. They also provided feedback on the scenarios under development by the SAIC staff.

Ten scenarios were developed which described potential applications of existing or emerging technologies and aspects of technologies. These technologies show promise for facilitating the access of individuals with sensory impairments to media and communications. Five scenarios were developed on technologies for the visually impaired, four for the hearing impaired and one that applies to both groups.

Visual Impairment

The first scenario applicable to the visual impairments category, Braille Devices to Allow Media Access, concluded that a major technology shift is required to design a full page Braille display to meet the media access needs of persons with vision impairments.

The second scenario, Input/Output Devices for Computer and Electronic Book Access, found primary solution strategies involve providing mechanisms to connect alternative display or display

Technologies for Visual Impairments

1. Braille Devices
2. Input/Output Devices
3. Visible Light Spectrum Manipulation
4. Flat Panel Displays
5. Descriptive Video

Technologies for Hearing Impairments

1. Adaptive Modems
2. Telecommunications Systems
3. Voice Recognition Systems
4. Video Teleconferencing

Technologies for Visual and/or Hearing Impairments

1. Portable Power Systems

translator devices to computers and provide alternatives to display-based input. Computer input device technologies attempt to solve the problems of mouse control and screen navigation in the absence of visual feedback or hand/eye coordination through such devices as touch-sensitive pads and charge-coupled-device (CCD) cameras. Computer output device technologies will attempt to provide alternative non-visual displays through voice synthesizers, Braille, or enhanced images through head-up displays adapted for the visually impaired.

Visible Light Spectrum Manipulation to Allow Media Access for Persons with Selective Vision was the third scenario. It established that computer access technology should be exploited to improve access to printed media and that night vision and image enhancement equipment could be adapted for use by persons with selective vision.

Flat-Panel Terminal Displays Used with Page Scanners, the fourth scenario, suggested that several key technologies can be utilized to provide visually impaired persons with access to flat-panel displays. These technologies include hand-held or flatbed scanners with appropriate optical character reader software, and a speech synthesizer package for voice output capability.

The fifth scenario, Descriptive Video (DV) for TV Access, found that the major commercial TV networks believe that DV is the right thing to do but are unable or unwilling to invest millions of dollars to produce and distribute an extra audio track for programs without some assurance that it will attract a large number of viewing households. Based on other studies it does not appear that there are any serious technical or regulatory obstacles. A limited number of DV series are broadcast at this time, most frequently on PBS; however, commercial viability is questionable without subsidies or a substantial non-visually impaired audience.

Hearing Impairment

The first scenario addressing persons with hearing impairments, Adaptive Modems and TDD Access, found no one addressing technology that combines access to telephone devices for the deaf (TDDs) with access to high-quality 9600 bit-per-second (bps) data transmission over standard phone lines for today's PC modem or FAX user. Instead, persons with severe hearing impairments must use outmoded Baudot TDD modems, which do not provide access to telecommunications services, such as person-to-person communications (except through other TDDs or relay services), electronic mail and database retrieval systems.

Telecommunications System Access, the second scenario addressing persons with hearing impairments, discusses limitations to accessing specific parts of the telecommunications network, including non-TDD-equipped individuals, voice mail, automated attendant systems, and Public Switch Exchanges. The technology areas which could benefit the hearing impaired include voice recognition systems to provide input to TDD relay services to help eliminate operator assistance, technology to provide call progress information by recognizing audio signals (i.e., a busy signal), and technology for access to automatic message answering systems.

Voice Recognition for Personal and Media Access concluded that in order to satisfy the requirements of persons with hearing impairments for natural voice processing, advanced technology voice recognition systems are required that are speaker-independent and can translate speech to text in real time. Several systems were identified in the scenario that offer the promise of this capability within the next three to five years. There has been a paradime shift in speech recognition systems in the past two years based on work by the Defense Advanced Research Project Agency.

Video Teleconferencing/Data Compression for Persons with Hearing Impairments, the fourth scenario addressing persons with hearing impairments, found that, in principle, the video telephone is far more attractive than the TDD to many deaf persons for communication with someone who knows sign language. However, video telephones which send pictures and accompanying voice conversations have been useless for sign language since a whole sequence of signs would be blurred into a single picture. Picturephones can transmit a picture useful for signing but require bandwidths which are 200 to 300 times the bandwidth of a standard phone line. Current research in video compression may be the answer to developing products which could use existing telephone lines to communicate sign language for persons with hearing impairments.

The last scenario, Portable Power Systems, is applicable to technology users with any type of sensory impairment. It contains compiled data that would enable an equipment designer to choose appropriate batteries for devices that benefit the sensory impaired. It could also be used to help ensure appropriate choices of portable power sources in current and future research and development efforts which the Government could fund to benefit the sensory impaired.

Recommendations

It is recommended that OSEP and the U.S. Department of Education continue to fund programs in media access and initiate programs in advanced technology fields as discussed in this report. Cooperative programs are recommended with other Government agencies and departments. This cooperation is essential in exploiting technologies such as speech recognition, actuators for Braille devices and new input/output devices. The detailed program recommendations in this final report specify how such programs should be structured and what areas of advanced technology should be explored.

A follow-on study is recommended in another three to five years to examine advanced technologies for benefits to persons with sensory impairments. The results of a follow-on study would be the formulation of another five- to six-year plan, taking advantage of new technology. Therefore, with technology changing at such a rapid rate, OSEP's planning cycle will be knowledgeable of events and discoveries that could not have been predicted more than five years in advance.

1.0 INTRODUCTION

This final report covers the work performed by Science Applications International Corporation (SAIC) on Department of Education Contract Number HS90047001. The program title is "Examining Advanced Technologies for Benefits to Persons with Sensory Impairments" was conducted between October 1, 1990 and March 3, 1992 for the U.S. Department of Education Office of Special Education.

1.1 Report Structure

This report is broken up into six sections and appendices. The first section "Introduction" establishes the framework of the project, including the structure of the report and the concept of the study. Section 2, "Purpose and Objectives" states the purpose of the study and the objectives of the project. Section 3,"Technical Approach" covers the approach taken to execute the program including the organization and conduct of the panel of experts meetings, the approach to the database searches, the approach to data collection and site visits, and the approach to scenario development. Section 4, "Results and Findings" discusses the results and findings of the program as they relate to advanced technologies which can benefit persons with sensory impairments. This section will also address specific needs and applications which can benefit sensory impaired persons. Section 5, "Conclusions" addresses the scenarios and how they relate to persons with sensory impairments. An estimate of the significance of the scenarios to target audiences will be projected for 3 to 5 year and 5 to 10 year timeframes. Section 6, "Recommendations" makes recommendations about the need for Department of Education involvement or sponsorship of particular advanced technologies which can benefit persons with sensory impairments. Appendix A is the programs Conceptual Framework that guided program management and Appendix B is the information collection plan that established the methodology for information collection over the 19 months of program execution. Appendix C, "Ten Year Development Plan" provides a ten year advanced technology action plan to assist the Department of Education in developing priorities that meet the most urgent media access needs of the sensory impaired.

1.2 **Background**

To ensure advanced technology research and development meets the needs of persons with sensory impairments, a comprehensive program was initiated by the Department of Education, Office of Special Education Programs (OSEP), to define the needs and recommend research goals for future Department of Education research and development programs. The goal was to determine priorities for advanced research and development for advanced technology device development, and to foster legislation to encourage private research and development and integration of special functions into products to allow media access for persons with sensory impairments. The concept of the program was to insure future advanced technology programs include a focused effort to adapt second and third generation media access products to meet the access needs of by persons with sensory impairments. The Program was also intended to evaluate design to cost for special features such as close captioning, descriptive video, etc., and make recommendations to OSEPs.

The initial step for the Department of Education was to establish an advanced technology program under the OSEP to define the program goals and focus the effort on specific sensory impairments. The second step was to initiate the program "Examining Advanced Technologies for Benefits to Persons with Sensory Impairments," a study to examine advanced technologies in both the private and public (U.S. Government Laboratories) sectors. SAIC was awarded the contract in October 1990 to perform the study. This final report is presented in fulfillment of the terms and conditions of the contract.

2.0 PURPOSE AND OBJECTIVES

This section covers the purpose and objectives of the program.

2.1 Purpose

The purpose of the program was to perform a technical study of Advanced Technologies for Benefits to Persons with Sensory Impairments, examining both the public and private sectors. SAIC as a diversified company which understands the technologies being applied within industry, research engineering laboratories, the military, and academia was selected to undertake this study to assist the OSEP as a first step in beginning a comprehensive advanced technology program.

2.2 Objectives

The objectives of this technical study were to identify:

- advanced or emerging technologies that can have implications for access of individuals with sensory impairments to media such as films, video, TV, print, telecommunications devices, electronic correspondence, and innovative uses of current communications devices (FAX, computers, page scanners, etc.);
- applications and features of applications that may facilitate the access of individuals with specific disabilities to media communications as well as factors that may limit their access;
- adaptations that can facilitate access and minimize barriers;
- development and evaluation activities necessary to achieve those adaptations, and the number and type of groups benefiting from technology applications.

2.3 Program Tasks

There were eleven tasks invoked in the program contract. These tasks were:

Task 1:	Program Plan Review
Task 2:	Form Expert Panel
Task 3:	Develop Conceptual Framework

Task 4:	Prepare Information Collection Plan
Task 5:	Panel Meeting #1
Task 6:	Final Conceptual Framework and Information Collection Plan
Task 7:	Implement Information Collection Plan
Task 8	Panel Meeting #2
Task 9	Develop Scenarios
Task 10	Prepare Final Report
Task 11	Performance Measurement System

This report provides a comprehensive review of the work accomplished over the 19 months of this contract based on these 11 tasks.

3.0 TECHNICAL APPROACH

SAIC pursued a disciplined approach to organizing and implementing this study. This approach was based upon our original proposal submitted to the Department of Education on June 22, 1990, and the Conceptual Framework document developed under Task 3 and dated March 4, 1991. The Conceptual Framework Document integrated ideas, techniques, technologies, and system concepts from many diverse sources into a program to meet the needs of persons with sensory impairments. This approach represented a significant risk reduction measure to organizing and implementing the study. The program tasks were divided into three different categories: (1) program management, (2) expert advice and oversight and (3) study execution tasks as shown in Figure 3.1.

3.1 Program Management and Control Tasks

Program management and control tasks included program planning (Task 1) and study performance measurement (Task 11). Our approach was to use the program plan and management controls to manage program implementation. It was particularly important to manage risk in the planning tasks (Tasks 3 and 4) and in the selection of data from which to develop a limited number of high-quality scenarios. The Performance Measurement System was implemented at program award and briefed to the COTR. The system proved to be more than adequate for the program and provided up-to-date cost and hours reporting that allowed the principal investigator and Department of Education program manager to quickly ascertain program status. The Conceptual Framework Document (Appendix A) and the Information Collection Plan (ICP) (Appendix B) provided the program road map. This road map was followed and led to the development of all scenarios. In particular, the ICP resulted in an efficient information collection process. At no time did SAIC have any trouble obtaining information from industry or Government organizations. In fact, we obtained so much we had to limit it.

3.2 Expert Advice and Oversight Tasks

The expert advice and oversight tasks were centered on the Panel of Experts listed in Table 3.2-1. SAIC used the experts' advice to guide the technical efforts and provide feedback to management control as a way of controlling risk, measuring progress, and planning expenditures. The Panel of Experts met twice during the program, once in the 4th month and once in the 15th month of the contract. It should be noted that each member of the Panel of Experts spent a minimum of two to six hours reviewing scenarios and providing comments. SAIC's principal investigator and technical writers used the Panel's comments to target and revise scenarios. The Panel's work guided SAIC's technical staff and resulted in the high quality scenarios at Appendix D.

The Panel of Experts provided contacts at universities, industry and the Government that resulted in information on specific advanced technologies being incorporated into the scenarios.

Dr. Mike Kelly of the Defense Advanced Research Project Agency (DARPA) provided the latest information on speech and natural language processing and how it could benefit persons with sensor impairments. Dr. Lawrence Scadden of the EIA Foundation and Dr. Tim Cranmer of NFB provided contacts and information on Braille devices and techniques. Dr. Judy Harkins provided exceptional recommendations on Telecommunications Devices for the Deaf (TDD). All Panel members made significant contributions to the program.

3.3 Study Execution Tasks

The study execution tasks formed the core of this study effort. In Task 3 the Conceptual Framework to guide project activities was developed for review by the Panel of Experts at their first meeting. In Task 6, the comments from the Panel of Experts and the Associated Conceptual Framework to guide project activities was reworked to provide the final Conceptual Framework and Information Collection Plan (ICP). The ICP was implemented in Task 7 with 15 site visits to various manufacturers and trade shows. These

Table 3.2-1. Distinguished Panel of Experts

Name	Organization	Expertise
Government		
Mr. Chet Avery	Director, Division for the Blind and Visually Impaired, U.S. Department of Education	Visual Impairments
Mr. Ernest Hairston (Program Sponsor)	Chief, Captioning and Adaptation Branch U.S. Department of Education, Office of Special Education Programs	Hearing Impairments
Mr. Guy Hammer	U.S. Department of Defense, Office of Technology Applications	Rehabilitation Engineering
Dr. Richard Johnson	Director, Rehabilitation Services Administration, U.S. Department of Education	Hearing Impairments
Dr. Mike Kelly	DARPA, Director Electronics Manufacturing	Manufacturing Engineering
Ms. JoAnn McCann	U.S. Department of Education, Office of Special Education Programs	Rehabilitation Programs Administration
Mr. Tony Valetta	Director, STAMIS, U.S. Army	Communicational Engineering, Parent
Business		
Dr. Tim Cranmer	Director of Research, National Federation of the Blind	Visual Impairments
Mr. Charles Estes	Executive Director, The National Association of the Deaf	Hearing Impairments
Mr. Nelson Dew	President, Dewtronics	Rehabilitation Engineering
Dr. Clint Gibler	Technical Staff, AT&T Laboratory	Rehabilitation Engineering
Dr. Judy Harkins	Director, Gallaudet Research Institute	Hearing Impairments
Dr. Carl Jensema Dr. Corinne Jensema Ms. Margaret Hardy Ms. Tina Downing-Wilson	The Conference Center	Hearing/Visual Impairments
Dr. Harry Levitt	Distinguished Professor of Speech and Hearing Sciences, Graduate School of City Univ. of NY	Hearing Impairments
Dr. Lawrence A. Scadden	Director, Rehabilitation Engineering Center, Electronic Industries Foundation	Visual Impairments
Mr. Elliot Schreier	Director, National Technical Center, American Foundation for the Blind	Visual Impairments
Mr. Dan Winfield	Research Triangle Institute, NASA	Rehabilitation Engineering
SAIC		
Dr. Candy Anderson	Systems Engineer	Neural Networks
Mr. Dan Hinton	Principal Investigator	Hearing/Visual Impairments, Communications Engineering, Parent
Mr. Charles Connolly	Research Engineer	Electrical Engineering

included the Consumer Electronic Show and the National Home Health Care Show. Each visit was recorded in the monthly status reports.

In Task 9, ten scenarios were developed which described potential applications of existing technologies and aspects of technology. These technologies show promise for facilitating the access of individuals with sensory impairments to media and communications (see Table 3.3-1). Five scenarios were developed on technologies for visual impairments, four scenarios were developed on technologies for hearing impairments, and one was developed for visual and/or hearing impairments. The scenario development included ongoing discussions with the Panel of Experts regarding their specific subjects of expertise and sensory impairments.

Table 3.3-1. List of Scenarios

Technologies for Visual Impairments	
1	Braille Devices and Techniques to Allow Media Access
2	Input/Output Devices for Computer & Electronic Book Access
3	Visible Light Spectrum Manipulation to Allow Media Access for Persons with Selective Vision
4	Flat Panel Terminal Displays Used with Page Scanners
5	Descriptive Video for Television Access
Technologies for Hearing Impairments	
1	Adaptive Modems and TDD Access
2	Telecommunications System Access (Touch Tone Signaling Access)
3	Voice Recognition Systems for Personal and Media Access
4	Video Teleconferencing/Data Compression for Persons w/Hearing Impairments
Technologies for Visual and/or Hearing Impairments	
1	Portable Power Systems

SAIC's methodology for program task execution is presented in Figure 3.3-1, Program Conceptual Framework. The implementation of this methodology drew on inputs from management and expert advice task outputs. The planning and implementation tasks were reviewed and guided by the advice of the Panel of Experts and the Department of Education's COTR. The relationship established between the Panel of Experts and SAIC's technical staff is shown in Figure 3.3-1. The arrows indicate an interactive relationship with a logical flow of advice and information between the SAIC team and the Panel of Experts. As the SAIC staff formulated the ICP and identified technologies for review at the Panel of Experts meetings, the principal investigator called specific Panel members for advice on technologies and their applicability to specific sensory impairments. Program control exercised oversight as the time phased tasks were executed, leading to this Final Report of the findings of the study. The COTRs oversight was established through the formal contract. This relationship is based on the deliverables in Table 3.3-2. The deliverables provided continuous program monitoring and control by the Department of Education's OSEP COTR. In addition, copies of the deliverables provided the Panel of Experts with a basis for expert advice and oversight.

The last and final category (Task 10) required the preparation of this final report and the Ten-Year Development Plan. The organization of the final report was provided in the conceptual framework and approved by the program manager for the Department of Education. It includes the Executive Summary, Introduction, Purpose and Objectives, Technical Approach, Results and Findings, Conclusions, and Recommendations as well as a Ten-Year Development Plan.

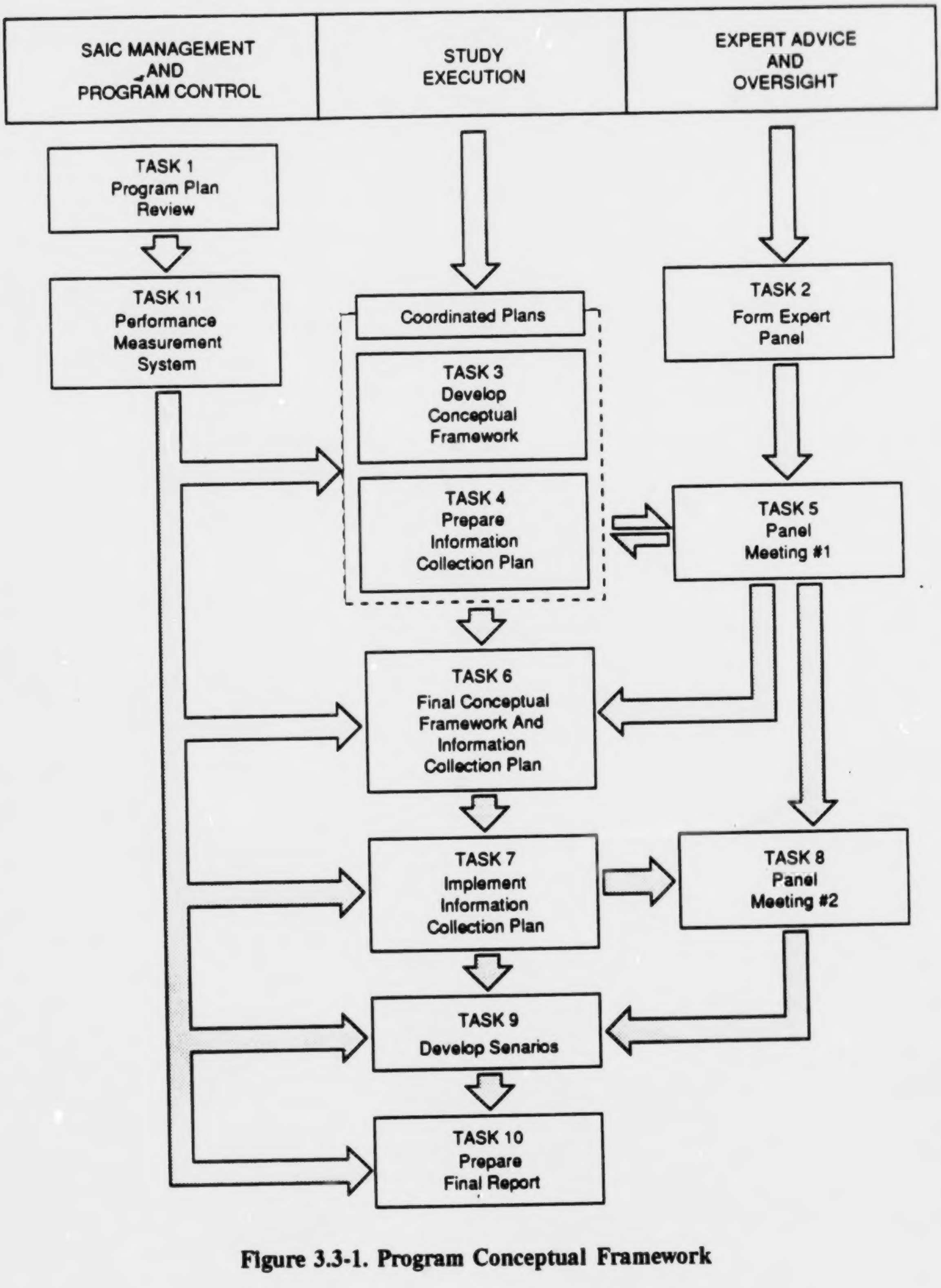

SAIC MANAGEMENT AND PROGRAM CONTROL
STUDY EXECUTION
EXPERT ADVICE AND OVERSIGHT
TASK 1 Program Plan Review
TASK 11 Performance Measurement System
Coordinated Plans
TASK 3 Develop Conceptual Framework
TASK 4 Prepare Information Collection Plan
TASK 2 Form Expert Panel
TASK 5 Panel Meeting #1
TASK 6 Final Conceptual Framework And Information Collection Plan
TASK 7 Implement Information Collection Plan
TASK 8 Panel Meeting #2
TASK 9 Develop Senarios
TASK 10 Prepare Final Report

Figure 3.3-1. Program Conceptual Framework

Table 3.3-2. List of Deliverables

ITEM	DATE DELIVERED	QUANTITY
List of Advisory Board Members (Task 2)	October 1, 1990	2
Administrative Reports (Task 11)	Monthly. See Figure 2.0-1	2
Draft of Conceptual Framework (Task 3)	December 10, 1990	2
Draft of Information Collection Plan (Task 4)	December 10, 1990	2
Final of Conceptual Framework and Information Collection Plan (Task 6)	February 26, 1991	2
List of Organizations Technologies (Task 6)	February 26, 1991	2
Case Report from Site Visits (when applicable) (Task 7)	October 25, 1991	2
Case Report from Site Visits (when applicable) (Task 7)	December 10, 1991	2
Draft of Scenarios (Task 9)	December 10, 1991	2
Draft Outline Final Report (Task 10)	March 10, 1992	2
Final Scenarios (Task 9)	March 10, 1992	2
Final Report and Ten-Year Development Plan (Task 10)	April 10, 1992	2

4.0 RESULTS AND FINDINGS

This section addresses the benefits of advanced technologies for persons with sensory impairments. Specific needs and applications will also be addressed. The complete impact of each advanced technology scenario, however can only be appreciated by reviewing the scenarios. There is just too much information in each scenario to cover it in detail in the final report. The scenarios are found at Appendix D.

4.1 Technologies for Visual Impairments

There were five scenarios developed for the category of technologies which might be beneficial to persons with visual impairments. The five scenarios were: 1) Braille Devices and Techniques to Allow Media Access, 2) Input/Output Devices for Computer and Electronic Book Access, 3) Visible Light Spectrum Manipulation to Allow Media Access for Persons with Selective Vision, 4) Flat Panel Terminal Displays Used with Page Scanners, and 5) Descriptive Video for Television Access. The results and findings of these scenarios are as follows:

4.1.1 Braille Devices and Techniques to Allow Media Access

Persons with severe visual impairments have limited real time access to computer information because existing Braille output devices are expensive and can only display 20-80 characters at a time. In the U.S., voice synthesis devices are used by more visually impaired Americans than paperless Braille devices due to their lower cost. Paperless Braille displays are more common in Europe, where the Government subsidizes devices for persons with impairments. Affordable paperless Braille is needed because voice synthesis does not allow the user to quickly review material as it appears on the monitor or printed page, including its format and structure. This makes it difficult to scan through text files and look for headings or jump from paragraph to paragraph. Thus, larger Braille displays are needed to allow persons with vision impairments text access capability equivalent to that of sighted persons.

Approximately 100,000 Americans with vision impairments use Braille for written communications. According to the 1988 National Health Interview Survey, 600,000 Americans between the ages of 18 and 69 have blindness or visual impairments severe enough to limit their employment opportunities. These two numbers provide an indication of the size of the population who could potentially benefit from Braille literacy.

The Department of Education and its predecessor, HEW, have funded Braille device research and development over the past 20 years. With the advent of personal computers the Government began to fund research and development of computer Braille output devices such as the TeleBrailler, and MicroBrailler. Currently, the development of Braille capability is a stated research priority of the Department of Education. The Americans with Disabilities Act established the objective of providing persons with disabilities access to physical and electronic facilities and media. The Technology-Related Assistance for Individuals with Disabilities Act of 1988 provided for technology access to persons with disabilities. These Acts cover the ability of persons with vision impairments to obtain access to printed media. Braille technology among others can provide persons with vision impairments with the opportunity to achieve literacy, participate in and contribute more fully to activities in their home, school and work environments, interact more fully with other people with or without sensory impairments, and engage in activities that are taken for granted by individuals who do not have severe sensory impairments.

Advanced Braille technology offers persons with visual impairments the potential for dramatic improvements in access to books and periodicals stored in computer-readable or scanned form. It is often desirable to skim text for relevant information, whether that text is a computer display, magazine or newspaper article, or book. A multiple line paperless Braille display would offer tremendous improvements in skimming speed and effectiveness over existing single line Braille displays and would be more convenient and potentially faster than existing Braille embossers for many applications. It would grant the person with vision impairments the opportunity to do research and academic study more efficiently, reading and rereading information with less effort and less paper output required.

Advanced Braille technologies contain two major approaches to producing paperless Braille. The simplest approach is to apply constant power to keep each dot raised or lowered, but many of the technologies used to move dots require a substantial amount of power to raise or lower the dots. Many paperless Braille displays raise or lower dots and then lock them into position until another page is displayed. The problem with locking mechanisms is that they increase mechanical complexity, thereby lowering the reliability and accuracy.

A major technology shift is required to design a full-page Braille display to meet the media access needs of persons with vision impairments. This new technology shift would incorporate advanced materials and computer control technologies. Advanced materials and manufacturing technology may make it possible to implement a display with several lines of Braille.

One current technology that could facilitate the implementation of full page paperless Braille is large array controllers for liquid crystal displays (LCDs). For several years, Smith-Kettlewell has been working on a proprietary electromagnetic Braille cell technology funded by the Department of Education's National Institute for Disability and Rehabilitation Research (NIDRR). Blazie Engineering has been working on a pneumatic display that uses puffs of air to move tiny bearings supporting Braille dots. Recent advances in sequential soft copy Braille displays have also been made.

Future developments for paperless Braille are impossible to predict with certainty because, though completely new technologies are seldom discovered, old ones are constantly revitalized by new computer capabilities, materials, and manufacturing processes. Sometimes older technologies suddenly become practical due to material or other technology breakthroughs. Three technologies which were suggested are: magnetostriction, electrorheological (ER) fluids, and polymer gels.

Magnetostriction is the property of some alloys that cause them to forcefully expand in a strong electromagnetic field. It does not appear to be a cost competitive technology for Braille cells , though this may change with future breakthroughs in material technologies.

Electrorheological (ER) fluids thicken when a strong electric field is applied to them on the order of 2000 volts per millimeter. This allows hydraulic actuators to be constructed. Since ER fluids stop flowing while in a strong electric field, they can selectively apply pressure to actuators. Three problems are likely to arise with ER fluid based Braille cells: the use of a liquid, difficulty modularizing a system with fluid lines, and fluid pump power and noise. ER fluids may be feasible for Braille cell development in the immediate future.

Polymer gels are another promising technology for full-page Braille displays. These gels collapse when exposed to intense light. Under the proper conditions, gels can be induced to reversibly release a large portion of their liquid content. Gel reaction times below a second require strands significantly thinner than a human hair. If one or more lasers were scanned over a Braille dot array with a mirror, gel technology might make it possible to implement a reliable full page Braille display with reasonable size, weight, power, and cost. However, this would require considerable development effort.

Smart materials combine sensors and actuators to react to special situations. They might be able to provide high reliability with imperfect locking mechanisms by verifying that a dot has been raised or lowered.

4.1.2 Input/Output Devices for Computer and Electronic Book Access

As computer displays become more visually complex, new strategies are needed to augment the standard approaches to providing persons with vision impairments access to the information being displayed. Problems associated with computer input devices deal with finding or identifying keys and controls on the keyboard and the problem of mouse-driven control. Visually impaired individuals have difficulty using perfectly flat membrane keyboards, since they cannot find the keys even if they have memorized their positions and functions. They also have difficulty in locating keys on large keyboards without tactile landmarks. Visually impaired individuals cannot use a mouse because they cannot monitor the mouse cursor's continually changing position as they move the mouse.

Problems with computer output devices deal with display and voice output. Some people with visual impairments cannot see lettering and symbols on keyboard equipment or screen because it is too small or too low contrast. They also need electronic access to information displayed on the screen in order to use non-visual display substitutes. Problems with computer input and output devices will become more severe in the future as computer systems move toward a more graphical approach to entering and displaying information.

Several microcomputer-based technologies have impacted the way visually impaired people access information from computers and electronic books. Computer access for visually impaired individuals is mostly limited by the cost, user interface and field of view of available displays, although newer display-based input systems (e.g., mice, touchscreens) may also pose problems. The visually impaired population includes individuals who have failing vision and individuals with partial vision, as well as those who are blind. The primary solution strategies involve providing a mechanism to connect alternate display or display translator devices to the computer, and providing alternatives to display-based input.

Technologies which are associated with computer input devices include Braille keyboards and optical character recognition (OCR). OCR is covered in a separate scenario titled, "Flat Panel Terminal Displays Used with Page Scanners."

Current technology output devices include: Braille output systems, speech synthesis, and large-print processing. Braille output systems are covered in a separate scenario titled, "Braille Devices and Techniques to Allow Media Access." Synthesized speech is one of the most powerful and least expensive access devices for the blind. Also available are many screen reader software packages designed to direct keyboard input and screen text output directly to a voice synthesizer.

Large-print processing is a valuable access medium for the visually impaired. Individuals with low vision may have difficulty reading the screen because the characters (text), or images are too small. In addition, they may have difficulty seeing the screen due to glare or distance. The two basic methods to add large print to an existing personal computer are to connect a hardware-based large-print processor or load a software package that remaps the characters of the video display to increase the size of the characters displayed. Software-based systems support a variety of computer functions such as word processing, graphics utilities, printer utilities, and Braille word processing.

Advanced media access technology offers the potential for dramatic improvements in information access for persons with vision impairments permitting direct access from existing and future computer-based information systems such as:

- Databases
- Electronic mail systems
- Bulletin board systems
- Mail order systems
- Books and articles.
- Screen graphics

Several new technologies are emerging which will greatly improve graphical computer interfaces for the visually impaired. Computer input device technologies will attempt to solve the problems of mouse control and screen navigation in the absence of visual feedback for hand/eye coordination. Three such emerging technologies are the "UnMouse," handwriting-recognition systems, and charge-coupled device (CCD) cameras. Speech recognition systems are also being pursued as computer input devices for the visually impaired. Computer output device technologies will attempt to provide alternate non-visual displays through voice synthesizers and touch screens, Braille, or enhanced images through head-up displays (HUDs).

Most advanced I/O device technology is relatively new and thus prices are high. As competition increases, the cost of input/output device technology is expected to decrease as with other computer-related equipment. For example, as the second and third generation voice recognition products begin to appear, the cost of the technology will be driven down by market forces and microelectronic implementations of voice recognition hardware.

Adapting certain advanced technologies for the purpose of enhanced computer access for the visually impaired may require a substantial investment that may not be practical for manufacturers without Government assistance or sponsorship for the initial research and development phases. The reason for this is that the sensory impaired market is small, which makes it difficult to recover development costs within a production run without passing the full cost onto a small number of consumers. The first applications for relatively small populations are therefore usually systems adapted from mass market devices. With a systematic approach to developing interfaces for applications for persons with visual impairments, the Department of Education can help reduce the cost of advanced input/output device technology to meet the needs of persons with visual impairments. This is possible because much of the research and development cost do not have to be amortized over the initial production runs.

The heaviest Government involvement in advanced input/output technology has been in the area of voice recognition. The U.S. Government involvement in voice recognition systems has been broad and includes National Security, Transportation, Commerce and Education applications. To date the most significant advanced technology effort is being conducted by DARPA's Information Science and Technology Office.

The Department of Education has maintained a large research program through both Grant and Small Business Innovative Research (SBIR) Programs for the past 30 years. The Department of Education programs provide the research and development platform essential to meet the computer input/output needs of persons with visual impairments. Without these programs to initiate new devices and probe new technologies, persons with sensory impairments would be denied access to advanced computer technologies.

The U.S. National Aeronautics and Space Administration (NASA) funds research and development efforts for devices for the handicapped through SBIR and university innovative grant programs.

4.1.3 Visible Light Spectrum Manipulation to Allow Media Access for Persons with Selective Vision

Visual impairment is the second major cause of disability in the United States. Nearly every American who lives to a normal life span eventually will be numbered among those who are considered "visually handicapped." Every man or woman over forty years of age require substantially more light when reading than their school-aged son or daughter. For many there is a constant deterioration of eyesight past middle age. Approximately fifty percent of us, when we have passed the age of sixty, will have a detectable degree of cataract development, and once we are into our sixties, most of us will have significant difficulty in determining small details. Virtually all who reach the age of eighty or beyond will experience major deterioration in vision due to either disease or the inroads of maturity.

Approximately 2.5 million Americans, many over 65, suffer from low vision. Birth defects, injuries, and aging can cause low vision, but most cases are due to eye conditions that affect the retina, including:

- Macular Degeneration
- Diabetic Retinopathy

- Glaucoma
- Retinitis Pigmentosa

Cataracts (the clouding of the lens), cornea infections, and detachment of the retina also can cause low vision. In addition to loss of central or side vision, low vision patients may lose their color eyesight, have difficulty adapting to bright and dim light, and suffer diminished focusing power.

Table 4.1.3-1 depicts some of the current products providing alternate displays which are available to persons with selective vision impairments to assist them in gaining media access.

Table 4.1.3-1. Alternate Display Systems Usable With All Software

Product Name	Vendor	Computer	Cost	Selective Vision Application
Advantage	Telesensory Systems Inc.	IBM	$31.95	Possible to have a positive or negative image on either half of the split screen.
Anti-Glare Magnification Screen	Sher-Mark Products, Inc.	Apple	$89.95	Polarizing filter is used to reduce glare and improve contrast.
CloseView	Apple Computer, Inc.	Apple	N/A	Screen display can be changed from black on white to white on black.
inLARGE	Berkeley System Design	Apple	$95.00	Both black-on-white and white-on-black displays are possible.
Large Print Display Processor	VTEK	Apple, IBM	N/A	Screen image can be positive or negative.
PC Lens	Arts Computer Products, Inc.	IBM	$690.00	Color options
Spy Graf	LS & S Group	IBM	$295.00	Color options
Zoomer	Kinetic Designs, Inc.	IBM	$89.00	Color options

Computer access technology should be exploited to increase the ability of persons with vision impairments to obtain access to printed media.

Advanced night-vision equipment now under development by the U.S. military may soon find a home in a host of commercial applications, thanks to the Army's decision to declassify its uncooled thermal imaging sensor technology. The new devices could be utilized by the visually impaired as well as the normal population in automobiles as "vision enhancers" for nighttime drivers. Japanese, German and US automakers have toyed with the idea of automotive night vision devices for some time, but have been deterred by extremely high costs. For persons with retinitis pigmentosa or other disorders, smaller and lighter devices will be possible for use at night or in the daytime for setting the contrast of objects.

Using imaging techniques developed for space exploration, NASA will develop a device designed to improve the eyesight of some 2.5 million Americans who suffer from low vision, a condition that cannot be corrected medically, surgically, or with prescription eyeglasses. The invention, called the Low Vision Enhancement System, will employ digital image processing technology. Experimenters will apply such processing techniques as magnification, spatial distortion, and contrast adjustment to compensate for blind spots in the patient's visual field.

The Low Vision Enhancement System should benefit patients with loss of central vision, the part of vision normally used for reading. These patients may have macular degeneration associated with aging, or diabetic retinopathy, in which diabetes causes swelling and leakage of fluid in the center of the retina. It also could help patients with impaired side vision due to eye diseases such as retinitis pigmentosa.

The Department of Education has conducted both Grant and SBIR Programs in this technology area for the past 30 years. These Department of Education programs are essential elements in meeting the computer input/output needs of persons with visual impairments. Without these programs, persons with sensory impairments would be denied access to advanced computer technologies.

4.1.4 Flat Panel Terminal Displays Used with Page Scanners

As computers become more visually complex, new strategies are needed to augment the standard approaches to providing persons with vision impairments with access to the information being displayed. The problem associated with computer input deals with character readers for accessing dynamic LED and LCD flat panel terminal displays. Computer output could be through voice synthesis, Braille, etc. This scenario points out key technologies which can be utilized to provide access to flat-panel displays:

- The user scans a text-based document from a flat panel terminal display into a PC using a hand-held or flatbed scanner.
- OCR software running on the PC "recognizes" bit-mapped characters in documents generated by the computer terminal.
- Some packages must be manually "trained" by the users to read new text. Other packages read any type automatically.
- Once the OCR software recognizes the bit-mapped characters, it translates them into a variety of text file formats, including ASCII and formats used by specific word processing programs. The files can then be called up from within word-processing or desktop publishing applications.
- A speech synthesizer package provides voice output capability.

Some observers think that the new flat-CRT technology is a viable way of producing CRT displays that could compete with LCDs for use in future laptop computers and other flat-panel applications. Such CRTs would be as thin and lightweight as LCDs, but brighter, less power hungry and cheaper to make. A vacuum-microelectronics display uses thousands of minute cone-shaped cathodes called microtips. They emit a stream of electrons that jump across a small vacuum gap toward a phosphor-coated anode to create images.

Current OCR products incorporate some form of automatic character recognition based on topological feature extraction, which consists of an algorithm that extracts salient features of each character and compares them to each other. Some of the programs use a form of matrix technology to aid in the recognition process. Because of the limitations of Charged Coupled Device (CCD) technology, most page scanners do not really capture a full 8 bits of usable information; electronic noise reduces the actual resolvability of the image to 7 or even 6 bits. Once the scanner creates the image, a high-speed direct-interface card transmits the image to the PC.

To capture color information, scanners make three passes, successively shining light through red, green, and blue filters. Eight bits of information are recorded for each color channel, providing up to 24-bit color.

Because of their limitations, hand-held scanners are not a suitable replacement for full-page desktop scanners. Most hand-held scanners can scan little more than a width of 4 inches in a single pass, although large images can be pieced together with multiple scans. Also, because most hand-held scanners are manually dragged across the image being scanned, image quality depends on how the user moves the scanner. The smoother and straighter the movement, the better the quality of the resulting scanned image.

Synthesized speech is one of the most powerful and least expensive access devices for the blind. Generally, a speech system consists of resident software that converts text into speech. When users optically read text, the system turns the letters into phonemes (the smallest units of sound), runs through a series of rules that tell it how to say the word, and outputs the word through the external speaker.

Several new technologies are emerging which will use OCR to enhance visually impaired person's access to flat panel terminal displays.

Synaptics, Inc. has developed an OCR system that it says is faster than existing solutions because image sensing and classification are performed in parallel on a single piece of silicon. The OCR chip packs an analog sensing array, two neural networks and a digital controller and extracts analog functionality from its digital circuitry. The new OCR chip operation is modeled on the human eye and ear that use digital circuitry to perform analog functions.

UMAX Technologies has developed a standalone OCR machine called ReadStation which combines a scanner, automatic document feeder, dedicated computer, and Caere Corporation's OmniPage OCR software. Printed or typewritten documents are fed into the ReadStation, converted to electronic form, and written as files to the built-in 3.5 inch disc drive. Word processing, spreadsheet, and database file formats such as WordPerfect, Lotus 1-2-3, and dBase are supported, and selectable using a control panel on the front of the unit. The unit can be connected to a PC via an RS-232C or RS-422 serial port interface for direct file transfer.

CCD cameras could be utilized as computer input devices. They would work like a scanner but be more portable. The CCD camera would use OCR software to read screens, books, or LCD displays, to name a few examples.

Handwriting recognition technology could also be tied in with OCR to enhance visually impaired persons' access to handwritten materials. This technology will allow a visually impaired person to be able to read mail, handwritten notes, etc. with little or no assistance. The enabling technology for this emerging market is the incorporation of neural-network techniques into a flexible object-oriented operating system. Pen-input computers are of little or no direct benefit to most of the severely visually impaired population, but their development has recently reawakened interest in handwriting recognition. System designers face several challenges, including: creating a system that can adapt to multiple writers handwriting, limiting the duration of system training, building a system that can recognize a wide enough range of characters, and allowing users to write naturally. The new technology will employ the following techniques to solve these challenges: examination of visual information; the handwritten text itself, analysis of data from the writing process, such as the sequence, timing, pressure and direction of pen strokes, and use of contextual data, such as predictable sequences of characters. Scanned handwritten text contains no time and pressure information, but recognizing it is otherwise analogous to recognizing text on a pen-input computer.

Voice synthesizer technology has seen rapid growth recently, especially in terms of improving the quality of the voice outputs. The focus is toward tailoring speech synthesis to the individual. By utilizing a smaller database of words based on a particular person's vocabulary, memory space and processing time can be reduced, thus allowing for the possibility of a higher quality of voice output.

4.1.5 Descriptive Video for TV Access

A visually impaired person in front of a television has limited access to information that is only presented visually. Described Video (DV) uses narration to describe the essential features of what is happening on the TV screen, omitting anything that is clear from the sound track alone. Video description can be anything from spontaneous comments to the scripted narration produced by any of several small private TV networks, up to the most carefully-developed scripted narration available.

Entertainment options for people with severe visual impairments are often limited. Many severe visual impairments make getting to places like movie theaters and playhouses difficult. Fifty-five percent of the severely visually impaired population are age 75 or older. Many visually impaired people, especially those who are elderly, have a fixed income. Most blind people are unemployed, and many people with visual impairments are underemployed. Described Television can provide a relatively inexpensive form of entertainment to these people, often the only entertainment available. The public broadcast system station, WGBH, estimated that 11.5 million people with visual impairments can benefit form DV, which is the approximate size of the visually impaired population, as shown in Table 4.1.5-1.

Table 4.1.5-1. Potential DVS Users by Level of Visual Impairment

Level of Visual Impairment	Estimated Population	Source	Date
Totally Blind--no or little sensitivity to light	0.05 million	American Foundation for the Blind (AFB)	1978
Legally Blind--acuity of 20/200 or worse in better eye with correction or a visual field of 20 degrees diameter or less	0.6 million	AFB	1986
Severely Visually Impaired--inability to read newsprint with corrective lenses	1.4 million	National Society to Prevent Blindness (NSPB)	1980
Severely Visually Impaired--inability to read newsprint with corrective lenses or, if under six years old, blind in both eyes, or having no useful vision in either eye	1.9-2.8 million	AFB	1986
Same as above; augmented by AFB's estimate of 500,000 institutionalized	2.4-3.3 million	AFB	1989
Visually Impaired--chronic or permanent defect resulting from disease, injury, or congenital malformation that results in trouble seeing in one or both eyes even when wearing glasses	8.4 million	National Center for Health Statistics (NCHS)	1988
Visually Impaired--same as above, includes color blindness, vision in only one eye, and other non-severe problems	12 million	WGBH testimony	1989

Producing and distributing described video demands careful planning and special equipment. The COSMOS Corporation study found that even the major networks feel that DV is the right thing to do, but they are unable or unwilling to invest millions of dollars to produce and distribute an extra audio track for programs without some assurance that it will attract about a million new viewing households.

The technologies capable of broadcasting described video (DV) on the Multichannel Television Sound (MTS) TV stereo system are summarized in Table 4.1.5-2. The present

Table 4.1.5-2. Technologies Capable of Broadcasting Described Video

Broadcast Technology	Pros	Cons
MTS Television Stereo System (Advantage: Sound is connected to picture.)		
Stereo Sum (Main Audio) Channel (15.2 kHz bandwidth)	All households can receive. All stations can transmit.	Reception not optional, so impractical for major networks.
Stereo Difference (Stereo) Channel (15 kHz bandwidth)	25% of households can receive, increasing. 48% of stations can transmit. Reception optional.	Only larger TVs can receive stereo now. Conflicts with stereo programs, and networks fear use would cause switching errors.
Second Audio Program (SAP) Channel (10 kHz bandwidth)	25% of households can receive, increasing. 20-48% of stations can transmit. Reception optional. 10% of stations use.	Only larger TVs can receive SAP now. Requires network to carry extra audio channel. Conflicts with second-language broadcasts, when available.
Professional (Pro) Channel (3.5 kHz bandwidth)	At most, 48% of stations can transmit. Reception optional. At most, 10% of stations use.	Practically no TVs can receive Pro now. Requires network to carry extra audio channel. Conflicts with intended use: station telemetry and cueing crews. Need signal processing to compensate for narrow bandwidth.
Special Modulation Techniques for TV (Advantage: Sound is connected to picture.)		
Vertical Blanking Interval (VBI) on TV Station (narrow bandwidth if only one VBI line used, easier with more than one line)	VBI can probably be routed through a major network's routing system and consoles without compromising on program timing.	VBI lines are in demand, but line(s) must be assigned to DVS. If VBI is used for final broadcast, need special receiver. Development required.
Advanced Speech Synthesis on Closed Captioning Channel	Narrow bandwidth required permits sharing closed captioning VBI line without conflict. Sending pronunciation cues could make sound better than text-to-speech. Sharing closed captioning VBI line guarantees channel availability. Can probably be recorded on most VCRs.	Speech quality must be investigated. Special decoder needed, based on decoders that will be required for closed captioning starting in 1993. Regulation required.
SCA on TV Station (narrow bandwidth)	Not widely used.	Need special receiver. Development required. Probably not technically feasible on station already using all MTS channels. Requires network to carry extra audio channel.
Spread Spectrum on TV Station	Used successfully in U.K. for high-fidelity sound (NICAM system).	Need special receiver. Development required. Regulation required. Requires network to carry extra audio channel.

Table 4.1.5-2. Technologies Capable of Broadcasting DVS (Continued)

Broadcast Technology	Pros	Cons
Radio Modulation Techniques (Disadvantage: Sound is not connected to picture.)		
Main Channel of FM or AM Radio Station (15 kHz bandwidth or 5-10 kHz bandwidth)	Accessible virtually anywhere by anyone (for example, in cars). May attract general audience even without picture.	Air time is expensive. Simulcast requires network to carry extra audio channel or synchronize tape.
SCA on FM Radio Station (5 kHz bandwidth)	Slots increasingly available. Two SCA channels per radio station.	Need special receiver. SCAs may be in higher demand than SAP channel. Simulcast requires network to carry extra audio channel or synchronize tape.
Radio Reading Services (which are SCAs) (5 kHz bandwidth)	Print disabled have access.	Only print disabled have access. Limited number of Radio Reading stations. Simulcast requires network to carry extra audio channel or synchronize tape.

technologies are : the stereo sum (main audio) channel, the stereo difference (stereo) channel, the Second Audio Program (SAP) channel, AM and FM radio stations, and radio subcarriers (SCAs), such as Radio Reader Services. The advanced technologies that are discussed in the specific scenarios are: Vertical Blanking Interval (VBI), advanced speech synthesis over the closed captioning channel, synchronous audio tape (which is an issue for stations, not consumers), the Professional (Pro) channel, developing new TV audio channels, and Advanced Television (ATV).

Based on the Smith-Kettlewell and COSMOS reports, it appears that the technical and regulatory environments offer no serious obstacles to the provision of DV services. However commercial viability, whether the cost of providing the service will be offset by a sufficiently large number of users to justify the cost without subsidies, is questionable, depending on marketing as much as cost.

The cost of DV services fall into two primary categories: those incurred by the provider of the services and those incurred by the user. The costs to the provider include network equipment modifications or adaptations, adaptation of existing computer equipment to compose narration to fit the programs, labor costs for creating the narration, the costs of coordinating with production studios, and finally , the cost of upgrading local (affiliate) station equipment to enable them to broadcast DV.

In a low usage scenario for a network, two hours of broadcasting each week, the cost would be approximately $1,702,800 or $5500 per hour of programming over a 3 year period. Under the high usage scenario of 50 hours of programming per week, the cost would be approximately $37,560,000 or $4,800 per hour of programming amortized over 3 years. It was estimated that non-labor cuts for upgrading the affiliate stations for DV would be less that $64 per hour of programming for the broadcasting of two hours per week and less that $3 per hour of programming for broadcasting 50 hours per week over a three year period. Labor costs would depend upon station layout.

The cost to the consumer of receiving DV programming is limited to the cost of the receiver. Cost estimates range from $50 for a SAP radio to about $150 for a decoder similar to a TV stereo decoder.

Government involvement in descriptive video technology is currently in the form of three Department of Education programs. The OSEP is supporting DV on PBS, primarily through the SAP channel, but also through Radio Reader Services. OSEP is also supporting DV on videotape. NIDRR is sponsoring research on transmitting video description on the TV vertical blanking interval (VBI).

PBS is now broadcasting eight described series over the SAP channel on 58 TV stations, with 14 Radio Reader Services providing an alternative or backup for areas that do not have SAP capable stations. At least one small private cable network broadcasts classic movies with descriptions on its main audio channel. The major commercial TV networks do not provide video description.

4.2 Technologies for Hearing Impairments

There were four scenarios developed for this category. The four scenarios were: 1) adaptive modems and TDD access, 2) Telecommunications system access (touch tone signaling access), 3) voice recognition systems for personal and media access, 4) video teleconferencing/data compression for persons with hearing impairments.

4.2.1 Adaptive Modems and TDD Access

Advanced telecommunications modems have been developed to meet the American consumer's needs for high quality data transmission at 9600 bps over standard telephone lines. At this time, access to this new technology by persons with sensory impairments is not being addressed by Government or industry (i.e., management, researchers, or marketers). This could perpetuate a situation in which persons with sensory impairments who use Baudot TDD modems have little or no access to telecommunications services (i.e., person-to-person communications, electronic mail and database retrieval systems). Unless action is taken, this barrier could persist into the foreseeable future.

New advanced microchip modem technology offers a leap forward in design flexibility over existing modems. Advanced modem technology now makes it possible to implement TDD modem functions in advanced ASCII modems. This may be accomplished through software resident on the modem chips or on the host computer system. Either way, expensive hardware modifications are not needed because the advanced modem technology uses digital signal processors, programmed for the modem tone generation and detection functions previously accomplished using expensive hardware.

The Department of Education has funded TDD modem research and development over the past 20 years. With the advent of personal computers in 1975, they began to fund research and development of dual capable Baudot TDD and ASCII computer modems specifically targeted for persons with hearing impairments. Presently, the development of Baudot TDD and TDD-compatible ASCII modems is a stated research priority of the Department of Education.

TDD modem access is a priority because there are an estimated 400,000 Baudot TDDs in use in the United States, a country with over 30 million hearing impaired people. However, the computer modems being used for communication between computers, on bulletin boards, Government information retrieval systems, and home shopping networks, to name only a few, employ ASCII modems that cannot be used with Baudot TDDs. Access to computer modem technology via Baudot TDDs has been limited to specially designed modems, due to the implementation of modem tone detection and generation functions in hardware. The Department of Education funded several of these TDD/ASCII modems that have a maximum data rate of 1,200 to 2,400 bps in the ASCII mode and 45.5 bps in the Baudot TDD mode. As advanced modem technology is applied, it will be necessary to either develop new limited market modems to meet the ever changing market, incorporate TDD modem functions into all advanced modem technology, or develop standards that require adding ASCII capability to all new TDDs.

Within one year of the enactment of the Americans for Disability Act (ADA) on July 26, 1991), the Federal Communications Commission (FCC) must prescribe regulations for TDD relay services which:

a) Establish functional requirements and guidelines.
b) Establish minimum standards for service.
c) Require 24 hour per day operation.
d) Require users to pay no more than equivalent voice services.
e) Prohibit refusing calls or limiting length of calls.
f) Prohibit disclosure of relay call content or keeping records of content.
g) Prohibit operators from altering a relayed conversation.

The FCC must ensure that the regulations encourage the use of existing technology and do not discourage or impair the development of improved technology. Thus a bridge between Baudot and ASCII equipment is required.

Advanced technology modem chip sets implement all modem functions in software on the chip sets, so all the modem manufacturers will have to do is write the user interface software or proprietary system functions. This includes the screen display formats, routines to save files that are received, and routines to send information files to the modem. The advanced modem chips have the capability to distinguish between ASCII data and voice, although they do not yet have word recognition capability. The advanced modem chip sets can also be programmed to emulate any dual-tone modem, such as the Baudot TDD modem function. A simple emulation program may be included on the chip set, resident on the host computer, or downloaded from the host computer into the modem chips memory.

This advanced modem technology offers the potential for dramatic improvements in telecommunications access for persons with sensory impairments, using their existing Baudot TDD modems to access:

- other modem users
- databases
- electronic mail systems
- bulletin board systems
- mail order systems

It is critical to recognize, however, that these improvements only come if the new modems support Baudot TDD access. Until then, although advanced modems may be easier to retrofit for Baudot TDD, the vast majority of modems will still be a barrier to improved media access for the hearing impaired who do not have ASCII-capable modems.

The key is that services that serve a broad segment of the general population will be among the first to use advanced modem technology to serve a broad segment of the general population. Advanced modems are backward compatible with most other modems because the advanced modem chip sets are able to distinguish the various modem formats and automatically configure themselves for the appropriate mode of transmission. With a Baudot TDD mode added as part of an enhanced instruction set, or as an externally programmable feature, any person with a Baudot TDD modem could access the systems discussed above, given software was added to allow the information to be displayed in a Baudot TDD compatible mode.

An alternative approach to providing ASCII modem access for persons with hearing impairments would be to require all Baudot TDD devices built after a specified date to be ASCII modem-compatible at the user level. This approach would specify a time frame in which all Baudot TDD modems for the hearing impaired would be converted to ASCII modem capability.

The advantage to making all new Baudot TDDs ASCII-compatible is that persons with hearing impairments move up into the computer-compatible modem world with little or no impact on existing computer modems. However, the effect of making all ASCII modems sold TDD-compatible would also be minimal, except on occasions when it makes communication with TDDs possible.

The number of companies that are developing--or have developed--advanced modem chip sets is growing rapidly. All these modems feature 9,600 bps full duplex operation which is more than 200 times the Baudot TDD rate of 45.5 bps. It is really more than 400 times as fast considering TDDs are only half-duplex modems, capable of communications in one direction at a time. Full duplex allows both parties to send information simultaneously, which, for one-on-one conversation, is much more important than very high data rates. Advanced modems also feature error correction to minimize the effects of noisy phone lines, and they can also perform data compression. Compression can increase effective data rates by factors of up to three to four times, making these modems more than 1000 times as fast as a Baudot TDDs in some applications. However, applications that demand high data rates involve data transfer between computers, not interpersonal communications.

Advanced modem prices have been falling sharply in recent years. TDD prices have been much more stable. It is projected that current top-of-the-line modems will be less expensive than TDD's within three years. Adding Baudot TDD function to advanced modems is estimated to cost $20,000 for each modem product line (the cost of adding about 100 lines of code to a program). In practice, the cost of developing an inexpensive feature--like Baudot TDD capability--is generally absorbed in a short time. Maintaining the additional program lines to support Baudot TDD capability over the life cycle of a modem would add about a penny to the retail price of each modem. In short, the per unit cost associated with adding Baudot TDD capability to advanced technology digital signal processor based modems is small, but that cost provides broad access to hundreds of thousands of Americans.

Looking at the long term picture (5-15 years), this small cost also enables the deaf community to slowly transition from the outdated Baudot 45.5 baud standard and transition to the technology being employed within the consumer electronics market. this transition would take about five years.

Within five years most interactive computer services will use the advanced modems, including Government, industry and educational institutions. In addition, several million individuals will be using these modems nationwide. The earlier a Baudot TDD standard is developed and required in all advanced technology modems, the less costly it will be to persons with sensory impairments. This is because it would make Baudot TDD-capable modems a mainstream consumer product. The installed base of advanced modems will then ensure access via software. This upgrade promotes the Department of Education goals through the implementation of Baudot TDD capability in all modems.

The Department of Education funded two modem projects that are listed in the FY89 NIDRR Program Directory. Many early TDD modem developments were funded by NIDRR and OSEP. One project was entitled "Integrated, Intelligent, and Interactive Telecommunication System for Hearing/Speech Impaired Persons." This Phase II project was awarded to Integrated Microcomputer Systems, Inc., Rockville, Maryland, and featured TTY/TDD and ASCII compatibility, "remote signal control, direct connection to the telephone system, and text-to-speech voice announcer."

A Field-Initiated Research project, entitled "Deaf-Blind Computer Terminal Interface," was awarded to SAIC in Arlington, Virginia, for the development of an acoustical modem interface between the Telebrailler, Microbrailler, TDD, IBM-PC compatibles, and the Commodore 64C.

4.2.2 Telecommunications System Access

A hearing impaired individual is challenged when he/she attempts to access certain parts of the telecommunications network, including: non-TDD equipped individuals, voice mail, automated attendant system, and Public Switch Exchanges (PBX).

Persons with hearing impairments are challenged by the expanded use of dual tone multi-frequency (DTMF) applications that make certain tasks easier for hearing people. These challenges fall into four basic areas: general telephone access, PBX and operator intercept, touch tone signaling access, and call progress monitoring.

By the same token, several potential applications of the DTMF system may increase the telephone access of hearing impaired people. For example, many customer support lines are now automated by using DTMF signaling to let the caller indicate his/her needs based on voice questions. This system could easily be adapted for use by the hearing impaired by providing a Baudot detection capability, possibly coupled with an advanced TDD that has a multiline display for the text.

The DTMF signaling system and the call progress tone standards are the basic technologies associated with telecommunication systems. Applications associated with telecommunications system access for the hearing impaired are directly related to these technologies.

An example of data communications using DTMF is the IBM augmented phone service (by IBM Entry Systems Division). This plug-in board & software allows a deaf person to communicate with a hearing person via the IBM computer (without a TDD). The user can type a message on the computer keyboard and the system will send it out over the phone line as synthesized speech. The person called presses keys on the telephone to "spell" the reply (i.e., "BOY" is 269); the software decodes the tones into possible words which the user reads on the computer screen. This provides a technique for basic telecommunication between the hearing impaired and non-TDD equipped individuals.

Other devices that provide use of the DTMF capability for telecommunications access may be exploited. Call progress monitoring is currently provided in relatively few TDDs, but could be added at a low cost. The Freedom 415 TDD by Selective Technologies, Inc., TDD and TouchTalk Travelpro by ZiCom Technologies, Inc., have built-in call progress monitoring to indicate dial tone, ringing, busy signal, or a voice answering.

A recent Department of Education SBIR Program Request for Proposal, Department of Education SBIR Request for proposal (RFP) #91-024, discussed several subjects related to telecommunications system access by individuals with hearing impairments. The list included: line status monitoring, a modem add-on device (ASCII), an auto-detect switch for FAX, ASCII, and voice calls, and 911 system operator training.

A need exists for an inexpensive device to assist persons with hearing impairments in detecting/identifying important line status signals.

Given the current high cost and relative rarity of modems that are both ASCII- and Baudot-capable, adaptation or development of an add-on device to allow standard ASCII modems to communicate with Baudot TDDs is an important issue. Such a solution would provide an easy, affordable way to communicate with a TDD via a personal computer (PC) and a hearing impaired person may also use the same PC to communicate with computer bulletin boards or other services using ASCII. This would eliminate the need for have both a standalone TDD and an ASCII modem.

As indicated in the SBIR RFP, a need has developed to discriminate between voice, FAX, ASCII, and Baudot TDD calls. Currently, the technology to automatically switch between FAX, ASCII, and voice exists. The extension of this technology to recognize Baudot TDDs is possible by adapting/developing an add-on controller. This could automate telecommunications tasks which often require human interaction.

Another area of interest is training for 911 system operators. The training material would teach 911 operators how to handle emergency situations involving people who are deaf or hard of hearing.

At the University of Delaware's Rehabilitation Engineering Center on Augmentative Communication, work is underway to define an integrated workstation for deaf individuals. The concept is to bring several applications together in a unified system that offers the advantages of the constituent parts. A key element in this work is to identify the modes of telecommunications that can be effectively used by deaf individuals. Specific areas of interest include: telephone monitoring, touch-tone decoding and voice response.

Persons with hearing impairments can benefit from enhanced access to telecommunication systems in the following areas: updating or verifying information in a remote computer database; message forwarding systems; financial transaction systems; alarm systems; energy management systems; credit card verification systems; and mail order systems. Persons with hearing impairments will also benefit from enhanced access to cellular telephone media.

Voice recognition, which is the subject of a separate scenario entitled "Voice Recognition Systems for Personal and Media Access," could significantly enhance the access of hearing impaired people. The voice recognizer could convert speech into text via TDD display or computer monitor for reading by the caller. To simplify recognition of synthesized speech, a synthesized speech standard could be developed. This would improve access to voice mail systems, automated attendant systems, and other voice-based systems.

There are two types of voice answering systems being used by industry. The first is used for voice mail. When the system answers, the caller is asked to enter a mailbox number and then leave a voice message. This type of system sometimes requires a password. The second type of system is designed to direct the caller to the right type of assistance within an organization. For example, when someone calls an insurance company, the voice answering system would say: if you are on a touch tone phone press 1 for policy renewal; press 2 for policy information; press 3 for operator assistance and so on until all services were covered.

The question is how to provide access to these systems for persons with hearing impairments. The voice mail system is the most difficult since it assumes voice-to-voice contact with no TDD or computer modems. However, if the person with a hearing impairment knows there is a voice answering system, then by observing the TDD light they might know to dial a number for an operator for TDD assistance. The automatic menu system could also be handled the same way as the voice mail systems. The only real difference in the two systems is the type of message they are trying to convey and what happens after a selection is made.

4.2.3 Voice Recognition Systems for Personal and Media Access

Advanced voice recognition systems are being developed to meet the needs of Government and the American consumer, for high quality data entry (transcription) and machine control. At this time, access to single word voice recognition technology for persons with sensory and physical impairments is being addressed by Government and research institutions (i.e., management, researchers, and marketers). Persons with hearing impairments need a high quality speaker independent continuous speech recognition system to provide interpreter services for face-to-face, public address, mass media and telephone media access. Because the single word voice recognition systems require a pause between words, they are limited to approximately 60 words per minute maximum and are not practical for use as an interpreter system for persons with hearing impairments. The average speaker speaks at a rate of 150 to 200 words per minute without pauses between words. For services such as closed captioning for television news, the rate can be as high as 270 words per minute. Clearly, to meet the requirements of persons with hearing impairments for natural voice processing, advanced technology voice recognition systems are required that are speaker independent and can translate speech to text in real time.

Voice recognition systems technology encompasses everything from simple user-trained computer software and electronic circuits capable of recognizing a few single utterances to user-adaptable speaker-independent continuous speech systems capable of recognizing 1,000 to over 20,000 different words. Although the speaker-dependent systems have been on the market for over 10 years, the advanced technology speaker-adaptable continuous speech recognition systems are just beginning to make their appearance, and the speaker-independent continuous speech recognition systems are in research and development. These systems are expected to be available within 3 to 5 years for specific applications such as medical transcription.

The advanced technology voice recognition systems today are using new digital signal processor boards, statistical software and advanced acoustic microphone technologies to achieve speaker-adaptable, speaker-independent continuous speech recognition systems that can recognize words, and form them into sentences in real time. The scenario on speech recognition summarizes the current state of technology, where it is expected to be in three to five years, and how it could be applied to meet the needs of persons with hearing impairments.

Voice recognition systems are divided into two classes: feature-based and speech-trained. Feature-based systems explore spoken words to determine characteristics of the vectors (i.e., composition of the words and spectral content) and to determine what common invariant behavior they have. From these vectors, characteristics rules are formulated which can then be applied to the recognition process. In speech-trained systems, speech is used to train the system automatically. There are currently three methods for accomplishing the training: template matching, statistical modeling, and neural networks.

Template systems are generally applied to single speaker voice recognition systems, although by training the system using several speakers, some degree of speaker independence can be achieved.

Statistical modeling systems have been developed because sound spectrum sequence analysis is too complicated at this time to determine all of the rules necessary to identify certain utterances as words. Template matching is impractical because the variability of pronunciation is too great, and phoneme templates have not been successful. To overcome the limitations of these systems, statistical models have been developed to extract the statistical properties of sound. These models are based on extremely simple machine states. The form of these states is assumed, and then their parameters are statistically estimated using a large amount of speech data. Currently the Hidden Markov Model (HMM) has been the most widely used statistical model.

What makes the Markov model "hidden" is when it is applied to speech recognition. Given a sequence of sounds (vectors), the model includes enough information to determine the probability that those sounds correspond to a given sequence of states, representing a particular word. However, there is no way to "see" which sequence of states produced the sounds; that information is hidden. All that can be done is to find the probabilities that various sequences of states (words) produced the observed utterance, then pick the one with the highest probability.

Neural networks have been used for small vocabulary (1-100 word) speaker-independent applications. However, as the number of words increases, the training time and complexity of the networks increases and system performance decreases. Also, there is no known efficient search technique for finding the best scoring segmentation in continuous speech using neural networks alone. To overcome these problems, hybrid systems are being employed that take advantage of HMM search capabilities to find the "N" best matches and then employ segmental neural networks (SNNs) computational advantages to evaluate those matches.

Presently the Department of Education has no investment in voice recognition systems. However many of the goals and objectives of the Department of Education could be met with a high quality user-independent voice recognition system.

Potential telecommunications and media access improvements for persons with hearing impairments will come with the advent of speaker-independent continuous voice recognition systems as follows:

- face-to-face communications with the general public;
- telecommunications media access;
- communications media access (TV, recordings, radio, public address systems);
- interpreter services (education, business).

These improvements will require developing:

- a word recognition database for the type of programs to be real time captioned;
- a training methodology for voice transcribers;
- an interface between the voice recognition system and the existing closed caption hardware.

The U.S. Government involvement in voice recognition systems has been broad and includes National Security, Transportation, Commerce and Educational applications. To date the most significant unclassified advanced technology effort is being conducted by DARPA's Information Science and Technology Office. DARPA has fostered research and development in speech and natural language systems for over 20 years. The DARPA work has generated interest throughout the Government and the civilian community.

4.2.4 Video Teleconferencing/Data Compression for Persons with Hearing Impairments

Approximately two million Americans have hearing impairments severe enough to make speech unintelligible with a hearing aid. Of these, "about 200,000 were born deaf or became deaf before they learned a spoken language, about 410,000 became deaf before the age of 19 years, and most of the remainder became deaf in later life due to the aging process.

"Sign languages enable the deaf to communicate...with great facility, in contrast to the difficulty with which the deaf communicate with the hearing community by means of reading lips and facial expressions, and by means of written messages. Because it can be easily learned and greatly speeds communication, American Sign Language (ASL) is known to the majority of congenitally deaf adults regardless of their educational background."

Two devices that are providing telecommunication for the deaf are telephone devices for the deaf (TDDs) and the video telephone. TDDs permit a sender to type messages to a receiver who sees the characters displayed on a screen or produced on another TDD. Although TDDs are useful for communication between deaf and hearing people, they have a practical disadvantage in that communication is slow and effortful when compared with voice or ASL communication.

The video telephone is far more attractive than the TDD to many deaf persons for communication with someone who knows sign language. "Video telephones that were intended to send pictures accompanying voice conversations, it should be noted, have been useless for sign language. A whole sequence of signs would be blurred into a single picture. These phones were not designed for real-time updates.

"The American video telephone (Picturephone) and the British version (Viewphone) both transmit a picture of the sender to the reader by means of a television raster scan. Unfortunately, Picturephone and Viewphone require a communication bandwidth of [1 MHz, which is 200-300 times the bandwidth available from standard phone lines]. Their enormous bandwidth appetite not only makes them unsuitable for existing telephone transmission and switching facilities, but it makes the development of video telephone facilities economically unattractive." Current research seeks to utilize advanced technology in video compression to develop products which could use existing telephone channels to communicate ASL and finger spelling for persons with hearing impairments.

It takes many seconds to transmit a clear detailed color picture, with accurate shading and textures, over a standard phone line even with image compression. Some pictures compress better than others, but using a standard phone line to transmit full-motion color video, at broadcast television quality, is presently beyond the state of the art. It is difficult to predict whether video compression will clear that hurdle by the time phone lines with enough bandwidth for video become cost-effective for individual use, but neither is likely in the next three to five years. So far, the bandwidth of a standard phone line is too restrictive to transmit such high-quality video at high frame rates.

Transmitting sign language, however, does not require anywhere near that video quality. The human mind can compensate for considerable loss in image fidelity. That compensation may require extra concentration when reading sign language, but many people with hearing impairments would prefer signing over a phone line to typing, especially since the native language of many people who were born deaf is ASL. To them, English is a second language, and they are often more familiar with ASL than English. Extra effort to read sign over a phone line may be preferable to typing on a TDD because signing can be faster and more expressive than typing.

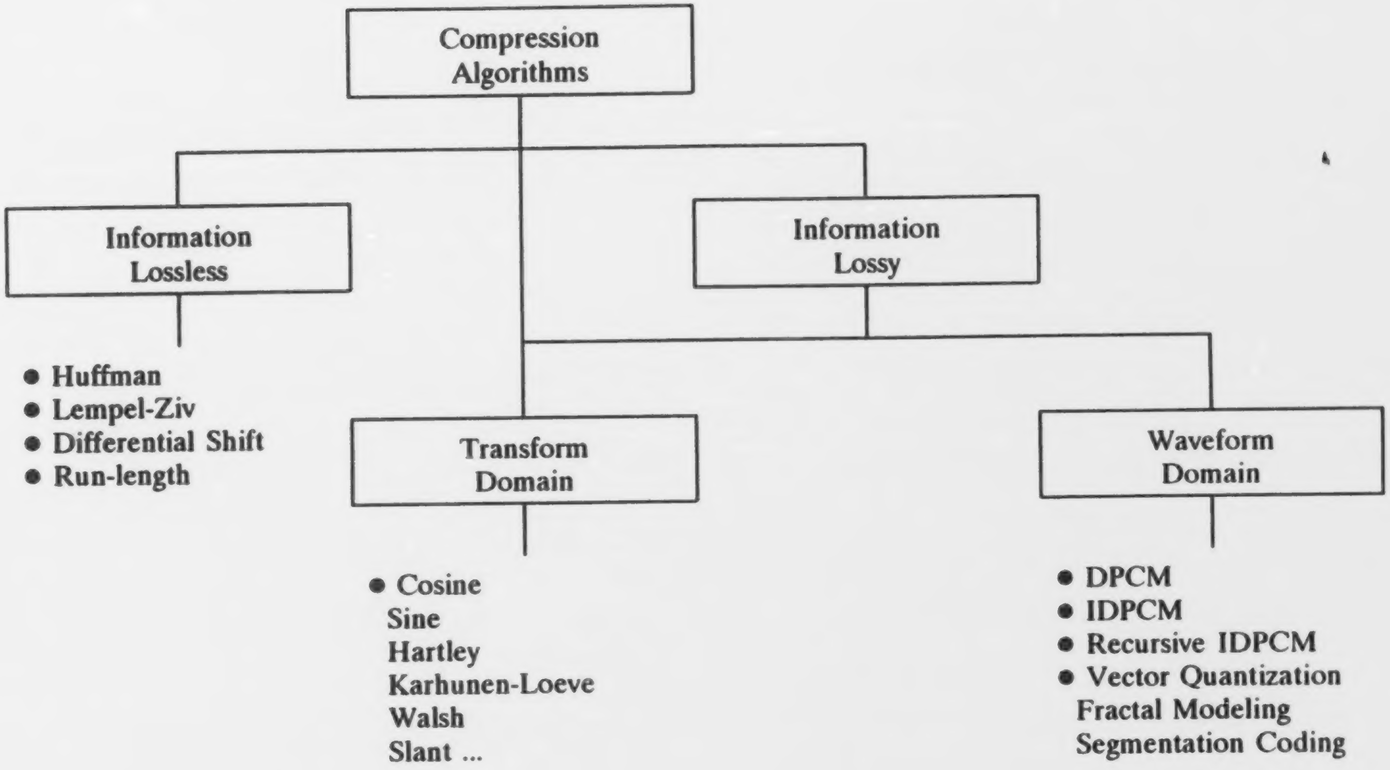
It should be noted that no technique for sending sign language over a standard phone line is in general use.

Common carriers, such as AT&T, MCI and Sprint, license video teleconferencing services through other companies, since divestiture does not permit them to deliver the services themselves. These services require the equivalent of many phone lines to transmit the video, however, and that makes them too costly for personal use. For business use, their cost-effectiveness would have to evaluated for individual cases, but they are probably too expensive for day-to-day use in most businesses.

These systems are optimized for high image quality, relative to frame rate. A video teleconferencing service optimized for much lower image quality at an adequate frame rate would be much more cost-effective for sign language, but only if there is enough demand for it.

When communication is to be over short distances, it may be economical to use simple video equipment and a cable to connect it. At Gallaudet University, for example, a low-cost crib monitoring camera and display are being used as an intercom between two offices. A simple video camera for finding out who is at the front door is another possible source of equipment for this type of innovative application. These approaches can be quite appropriate over short distances and should be publicized, but the cost of video-bandwidth cable becomes cost-prohibitive as distances increase, since it involves a per-foot cost plus an installation cost.

Many of the applications of digital video hinge on the use of image data compression, which means representing images in a more compact way to reduce the bandwidth required to send the images. Compression algorithms fall into two principal categories: information lossless and information lossy. Lossless compression means that in the absence of noise on the communication line, the original image can be reconstructed *exactly* at the receiver. Information lossy compression, on the other hand, means that some error is introduced by the compression process itself. The objective of image compression algorithm development is to minimize the visual impact of these errors. A taxonomy of popular compression schemes is shown in Figure 4.2.4-1.


Compression Algorithms
Information Lossless
Information Lossy
● Huffman
● Lempel-Ziv
● Differential Shift
● Run-length
Transform Domain
Waveform Domain
● Cosine
Sine
Hartley
Karhunen-Loeve
Walsh
Slant ...
● DPCM
● IDPCM
● Recursive IDPCM
● Vector Quantization
Fractal Modeling
Segmentation Coding

NOTE: Those with *'s are reviewed in this document.

Figure 4.2.4-1. Taxonomy of Compression Algorithms

Published results on the application of lossless algorithms to image data show that the compression ratios average about 2.2:1. Thus compression ratios of 500:1 or greater, needed to transmit sign language over a phone line, require incorporating a lossy algorithm. Transmitting sign language over computer networks, however, requires far less compression, sometimes requiring only reduced resolution; compression helps though.

NIDRR is currently funding research on transmitting sign language over standard telephone lines and over computer networks. That research is funded as two Field-Initiated Research Grants at the Department of Computer and Information Science, University of Delaware, Newark.

4.3 Technologies for Visual and/or Hearing Impairments

One scenario applies to both consumers with visual impairments and consumers with hearing impairments: Portable Power Systems. Today's electronic equipment is getting smaller and more complex. Portable power supplies that energize these electronic devices must be able to handle the discharge rates and also be economical. The data developed in this scenario will help the equipment designer choose a battery that is appropriate for a particular application that can benefit persons with sensory impairments. The Government will benefit from this study by gaining information to guide research and development efforts related to access for persons with visual and/or hearing impairments.

This report concentrated on miniature and portable equipment batteries. Typical uses of these are watches, calculators, medical devices, and small portable electronic equipment. Primary batteries are for one time use, non-rechargeable, and have been in use since 1866. They are generally low-priced, easy to produce and provide good performance. Secondary batteries can be recharged repeatedly during their lifetime. They are more expensive than primary batteries and also require a charger. The characteristics of the nine kinds of primary and secondary commercial batteries covered in the study are summarized in Table 4.3-1. They are listed in order of theoretical capacity.

Table 4.3-1. Commercial Batteries

Type	General characteristics	Package	Capacity (AH)[*]	Cell V	WH/lb	WH/in³	Operating temperature, °F (°C)	Shelf life (for 80% capacity)	Applications
	Primary Cells								
Zinc-air	Highest energy density on continuous discharge; excellent shelf life (unactivated); limited rate capability and shelf life when activated; flat discharge characteristic	Button Larger	1.15 8.5	1.4	140	19.0	32 - 122 (0 - 50)	5-10 years	For frequent use: Hearing aids, medical monitoring, pagers, communication equipment, data loggers, emergency lighting
Lithium-sulfur dioxide	High energy density; excellent high-rate and low-temperature performance; pressurized, hermetically sealed cell; excellent shelf life when activated; flat discharge characteristic	Cylindrical	0.7-19	3.0	125	7.1	-40 - 160 (-40 - 71)	5-10 years	High capacity, high-rate and extreme-temperature operation: Military and industrial applications
Lithium-manganese dioxide	High energy density; good rate capability and low-temperature performance; excellent shelf life; relatively flat discharge characteristic	Coin Cylindrical Bobbin	0.5 1.25 2.5	3.0	80 105 135	8.3 8.3 11.5	-4 - 140 (-20 - 60) -4 - 160 (-20 - 71)	5-10 years	General purpose applications requiring small high-capacity batteries: Watches, calculators, computers, memory backup, photoflash, camera motors
Alkaline	Popular, general-purpose premium battery; good low-temperature and high-rate performance; good shelf life; sloping discharge characteristic	Small Button Cylindrical	20	1.5	25 60	2.4 5.2	-4 - 130 (-20 - 54)	3-4 years	General purpose high-drain applications: Lighting, cameras, tape recorders
Zinc-carbon	Popular, common, low-cost primary battery; moderate shelf life; sloping discharge characteristic	Cylindrical	40	1.5	34	2.3	23 - 113 (-5 - 45)	1-2 years	General purpose applications: Lighting, radios, novelties
Mercuric oxide	High capacity per unit volume; good shelf life; flat discharge characteristic	Button Cylindrical Larger	0.04-13	1.35 or 1.40	50 55	7.3 7.4	15 - 130 (-9 - 54)	3-5 years	For steady output voltage: Hearing aids, medical and photographic equipment, communication devices, pagers, detection and emergency equipment
Silver oxide	High capacity per unit weight; good shelf life; flat discharge characteristic	Button	0.18	1.5	60	8.2	-4 - 130 (-20 - 54)	2-3 years	Small high-capacity batteries: Hearing aids, watches, photographic equipment, and special electronics
	Secondary Cells								
Nickel-cadmium	Sealed: Maintenance free; good high-rate, low-temperature performance; excellent shelf life	Button Cylindrical	0.5 10	1.2	10.6 45	1.0 1.4	-40 - 113 (-40 - 45)	3-6 months	Portable tools, appliances, televisions, computers, memory backup, electronics
Lead-acid	Sealed: Maintenance free; low cost; good float capability	Cylindrical Flat-Plate	2.5-30 50	2.0	13.5 15	1.5 1.3	-40 - 140 (-40 - 60)	6-12 months	Portable tools, appliances, televisions, computers, electronics

[*] Button cells are rated at the 500-1000-hour rate at 70°F (21°C), cylindrical cells at the 50-100-hour rate. The cutoff voltage is approximately 80% of the nominal voltage. Typical capabilities may be higher.

BEST COPY AVAILABLE

This report covered the Zinc-based batteries such as Zinc Carbon, Zinc Chloride, Alkaline Manganese, Aluminum/Magnesium Leclanche, Mercuric Oxide/Silver Oxide, Zinc Mercuric Oxide, Zinc Silver Oxide, and Zinc Air. It also covered the Lithium-based Solid Cathode, Lithium Sulphur Dioxide, and Lithium Liquid Cathodes. Because conventional dry batteries such as the zinc carbon have reached their technological peak and mercury/silver batteries do not meet the required power levels, Lithium-anode-based batteries have been researched. They have been found to yield energy densities up to 3 times that of the mercury- and alkaline-based batteries and volumetric densities of 50 to 100% higher.

The types of secondary batteries covered include Lead Acid, Nickel Cadmium, Nickel Hydride, Zinc Silver, and Conductive Polymers.

5.0 CONCLUSIONS

5.1 Technologies for Visual Impairments

5.1.1 Braille Devices and Techniques to Allow Media Access

Based on the outlook for affordable full-page Braille displays, two compromise approaches should be considered. First, smart materials are a way of increasing the reliability of existing mechanical locking mechanisms, although eliminating the need for locking mechanisms might be better in the long run. Second, further investigation into the effectiveness of a sliding one-line display is justified by the lack of compelling evidence that a full-page Braille display is technologically feasible in the next 3 to 5 years. Research and development efforts should be devised to push advanced materials and manufacturing technologies for Braille beyond the laboratory stage.

5.1.2 Input/Output Devices for Computer and Electronic Book Access

Most of the advanced technologies for enhanced computer and electronic book access have had or will soon have first generation products on the market. For example, within the next one to two years, several user-independent continuous voice recognition systems are expected to be marketed based on the research sponsored by DARPA and private companies, such as, the American Telephone and Telegraph Corporation. Several advanced input/output technologies are expected to mature over the next five years to the point where they will provide computer control for persons with visual impairments. What is needed are comprehensive programs to apply the technologies to meet specific needs of persons with visual impairments. This will require that training programs be formulated and specific goals set to allow the input/output technologies to be adapted for use by persons with visual impairments.

5.1.3 Visible Light Spectrum Manipulation to Allow Media Access for Persons with Selective Vision

The Innovative Research Grants programs sponsored by NIDRR would be the optimum tool to encourage I/O device development, along with the SBIR program. However, at least one program using a contracts mechanism should be considered for advanced speech recognition and synthetic voice to encourage large businesses to enter the field of I/O device development for sensory impaired persons.

Advanced technologies for enhanced computer and electronic book access have had or will soon have first generation products on the market too. Many of the advanced light manipulation technologies will mature in the next three to five years, and if directed at the visually impaired, provide computer access. What is required is a comprehensive programs to apply the technologies to meet specific needs of persons with visual impairments, such as large character displays and night vision devices. The Small Business Innovative Grants Program (SBIR) would be an excellent vehicle to foster this research into specific applications because many of the companies developing the hardware and software employ less than 100 people.

5.1.4 Flat Panel Terminal Displays Used With Page Scanners

Advanced technologies for optical character recognition have matured and first generation products are on the market. Advanced input/output technologies are mature to the point where they can provide computer control for persons with visual impairments. What is needed is a comprehensive program to apply the technologies to meet specific needs of persons with visual impairments. Specific goals must be set to allow the OCR technologies to be adapted for use by persons with visual impairments. The innovative grants process may be the best vehicle to encourage this work on application of OCR devices.

5.1.5 Descriptive Video for Television Access

At present, for transmitting descriptive video (DV) from network affiliate stations to homes, the "SAP channel is the only practical medium" that allows both audio and video to be sent on the same carrier. Unfortunately, the equipment and labor costs of adapting an entire network to transmit DV are presently high, due primarily to the cost and potential complexity of handling an extra audio channel at the network facility.

The most promising solution to the network facility problem is to distribute the extra audio channel over the vertical blanking interval (VBI) of the network's video signal. That way, the network facility remains intact, and the extra audio channel is inherently sent wherever the picture goes. Affiliate stations can then decode the VBI signal and impress the resulting audio onto the SAP channel at the facility where they normally add subcarriers. Advanced TV systems may be on the market as soon as 1993 or 1994, and they can incorporate an audio channel dedicated to DV if the FCC requires its inclusion. The Department of Education NIDRR is funding a grant on VBI equipment development. NIDRR should continue this effort.

5.2 Technologies for Hearing Impairment

5.2.1 Adaptive Modems and TDD Access

In the mid- to late-1990s, most computers will be user friendly and require no more computer literacy than today's Baudot TDDs. TDDs will also cost more than a much more flexible mass produced computer. This is already happening. By the year 2000, manufacturers will not be able to recover their costs when they try to sell new TDDs at prices competitive with computer equipment. Several hundred thousand people will still have traditional Baudot TDDs.

5.2.2 Telecommunications System Access

The future of advanced technologies for telecommunications access is telephone relay systems for the hearing impaired. These systems will become more and more automated. Automation will depend on ASCII capability and advance toward technology which makes the current ASCII-Baudot distinction transparent to the user. Eventually relay systems will switch to machine operated relays based primarily on ASCII. The basic idea is that the hearing impaired will no longer need human intervention to achieve access to telecommunications.

5.2.3 Voice Recognition Systems for Personal and Media Access

DARPA sponsors much of the most advanced unclassified research in speech recognition in the U.S. These speech recognition systems represent significant steps forward in user-dependent and continuous user-independent speech recognition systems that can be applied to the needs of persons with hearing impairments and physical handicaps. However, machines can improve the output to the person with hearing impairment by having the capability to examine the language structure to check for misinterpreted words and phrases. Therefore, some of the systems go beyond voice recognition systems and are speech and natural language processing systems.

Incorporating voice recognition capabilities into devices such as TDD phone relay, closed caption or interpreter services will require a substantial investment that may not be practical for manufacturers of voice recognition systems without Government assistance or sponsorship for the initial research and development phases. This is because the handicapped market is small, and it is difficult to recover development costs within a production run without passing the full cost on to a relatively small number of consumers. The first applications will therefore be systems adapted from mass market devices such as transcription systems for doctors.

Voice recognition technology is expected to mature over the next five years to the point where it will provide transcription, computer control, and interpreter services for persons with hearing impairments.

5.2.4 Video Teleconferencing/Data Compression for Persons with Hearing Impairments

Persons with hearing impairments can potentially benefit from advanced video compression technology because it can enable them to communicate with friends, relatives, and coworkers in sign language. This is an extremely important advance because hearing impaired persons, especially those who were born deaf, are often accustomed to communicating through sign language. Thus, the use of English for TDD communication is often difficult and uncomfortable.

Relay services for the deaf would also greatly enhance communication between the hearing and hearing impaired communities if they accommodated the use of sign language in addition to the use of TDDs. Conversation would be potentially much faster and more natural through the use of sign language.

Video services over phone lines could also make it possible to share the resource of sign language interpreters.

Video services may also benefit persons with hearing impairments who rely on lip-reading in combination with hearing speech. Likewise, cued speech, a technique developed at Gallaudet University for providing visual cues with speech for the hearing impaired listener, would benefit from providing video with speech.

Computer-generated cartoons, produced by a signal processing technique called edge detection, are currently favored as a potential technique for transmitting signs over a telephone line.

The image processing required for edge detection is expensive, but that cost will be brought down by the use of application-specific integrated circuits (ASICS) in the next few years. Edge detection has many other applications such as surveillance and robotics, so it is also of interest to the military, for law enforcement, for industrial applications, and eventually, for consumer applications.

The Department of Education should establish a major program to exploit video compression under a contract or grant to apply advanced video compression to specific applications for persons with hearing impairments.

The quality of sign language cartoons may also benefit from the use of anti-aliasing, which is a technique for smoothing the jagged, grainy lines of low-resolution images. Anti-aliasing can only help if the display is essentially better than the image it is displaying, but that may often be the case. A small display may be used to make a low-resolution image look better, but it may be worth considering the option of using a larger display with anti-aliasing, to reduce eye strain.

With the possible exception of edge detection, fractal compression is the most promising video compression technology for extremely high compression ratios, given that phone transmission of sign language has a high tolerance for selectively throwing out video information. Fractal compression is based on generating a mathematical representation of aspects of a picture, based on repeated patterns called fractals that often occur in nature (pine cones, for example, are fractals). Fractal compression can require a great deal of processing power for some applications, but it should be investigated for sending sign language over a phone line because it tends to produce hierarchical representations of images. Only certain image details are required to represent sign language intelligibly, and fractals may be helpful in selecting the right details.

Computer networks are typically designed to carry much more bandwidth than standard telephone lines. Often, computer networks go through the telephone system, but sizeable networks typically use specially installed digital lines, which would be capable of carrying many telephone conversations at once. Local area networks (LANs) also use cables that provide far more bandwidth than a standard telephone line can carry. That extra bandwidth makes it possible for many computer networks to carry sign-language conversations with far less image data compression than would be required over phone lines. In some cases, the frame rate and resolution can be reduced without the need for any other form of image data compression.

"The advantage of using the [computer network] over the sign language telephone is that you can see a real person rather than a line-drawn representation of the person. Certain nuances of signing can be more readily understood by seeing a real person versus a line drawing. Also, such restrictions as having to wear a dark solid top are eliminated by using the [computer network approach]." Of course the computer network approach also requires a computer network that has at least some extra bandwidth. Not every office has a computer network, and an overloaded network may be inadequate to support the addition of video for sign language.

Existing technologies for commercial video teleconferencing are expensive when applied to sign language transmission, but this is largely due to their being optimized for high picture quality rather than high frame rate. This is not so much a technical issue as a cost/demand issue, since developing and fielding a low-image-quality high-frame-rate video teleconferencing system has evidently not been a high priority of companies that offer video teleconferencing services.

Eventually the Integrated Services Digital Network (ISDN) will probably cover all homes in the U.S., providing video bandwidth over ISDN lines at a cost that will make their use popular in homes. That would probably solve, or lay the groundwork for solving, most of the telephone access problems currently faced by the hearing impaired. However, universal availability of ISDN requires replacing the existing telephone network with wider-bandwidth transmission lines, such as fiber optics. Completing that upgrade in a few years would be cost-prohibitive.

ISDN is already available in some areas and some buildings, however, and increasing network bandwidth should cause wideband channels to slowly become more affordable.

Sign language relay systems and interpreter-sharing systems are still two or three years away, at least, but it is important for the Department of Education to anticipate their development and be prepared; otherwise, they will take much longer to develop. A study could be very helpful in this regard.

Last, but certainly not least, computer networks could start to play a role in transmission of sign language in a year or two, but only if a sufficient number of options are investigated and publicized. Efforts to that end are likely to result in real progress.

5.3 Technology for Visual and/or Hearing Impairments

No special barriers to portable power technology access are foreseen, other than cost and standardization. Batteries of at least flashlight size have significant differences between the few commonly used sizes and shapes. The standard sizes and ratings of smaller batteries used for hearing aids, watches, etc., fill catalog. If battery manufacturers could reach a compromise and produce fewer small standard batteries, it would encourage competition and lower prices, so all consumers would benefit, especially people with impairments that make them dependent on battery powered technologies.

6.0 RECOMMENDATIONS

6.1 Technologies for Visual Impairment

6.1.1 Braille Devices and Techniques to Allow Media Access

The Department of Education, as a first step in Braille cell development, should explore the possibility of a cooperative effort with both the U.S. National Aeronautics and Space Administration (NASA) and the Department of Defense (DOD) in the area of small actuators for use in full page Braille cells. Specifically, the Department of Education should institute a research and development program for small low power consumption actuators for use in Braille display devices, robotics and space applications. The work should start by having NASA and DOD sponsor the development of actuators taking Braille cell applications int consideration. The Department of Education could then sponsor a program to develop a single-line display and a program to develop a full-page multiple-line Braille display.

6.1.2 Input/Output Devices for Computer and Electronic Book Access

A review of the Grant and SBIR programs should be conducted over the next two years to determine the most promising input/output devices to allow computer access. This review should provide a comprehensive list of priorities for future grant and SBIR funding efforts especially in the areas of voice recognition, handwriting recognition, CCD cameras, speech synthesis, heads up displays, and Braille technology. Following this review, the Department of Education should establish a program to fund the most promising techniques over a three to five year period.

The Department of Education should establish a program to fund two or three devices covered in the input/output devices scenario into an advanced development phase. This will allow a few small businesses to implement input/output devices and help move the devices from the development phase to the production phase. This will ensure continued computer access for persons with vision impairments. Overall, the Grant and SBIR programs should be continued as structured to encourage the development of input/output devices for the visually impaired.

The most promising programs from the SBIR's should be recommended for Innovative Grant programs. Field Initiated Grant programs should continue to be pursued when deemed appropriate. Because most of the technologies involved in the input/output device scenario are being developed for other commercial applications, 3-4 years seems a reasonable time period for each program. The payoff at the end of 3-4 years is the empowerment of persons with visual impairments to allow them to use systems that allow them equal access to computers and electronic books as well as access to personal communications services.

6.1.3 Visible Light Spectrum Manipulation to Allow Media Access for Persons with Selective Vision

The Department of Education should begin the process of developing devices utilizing advanced light spectrum manipulation technology for computer and electronic book access for persons with visual impairments by developing several key technologies. SBIR grants should be initiated in the areas of infrared sensors, digital image processing, CCD cameras, and heads-up displays.

The most promising programs from the SBIR's should be recommended for Innovative Grant programs. Field Initiated Grant programs should continue to be pursued when deemed appropriate. Because most of the technologies involved in this scenario are being developed for other commercial applications, 3-4 years seems a reasonable time period for each program. The payoff at the end of 3-4 years is systems that allow equal access to computers, electronic books and personal communications services for persons with selective vision.

6.1.4 Flat Panel Terminal Displays Used with Page Scanners

The Department of Education should begin the process of developing advanced OCR technology devices for use by persons with visual impairments by encouraging integration of several key technologies. SBIR Grants should be initiated for experimenting with the use of CCD cameras and other scanning technologies with flat panel terminal displays, LED and LCD displays, eventually providing output through speech synthesis.

6.1.5 Descriptive Video for TV Access

The COSMOS study of the commercial viability of descriptive video (DV) concluded that supportive marketing conditions would be needed for DV to be produced and distributed by the commercial networks. Conditions would include both startup and production investments, legislation, and FCC regulations.

An important role of the Department of Education would be to ensure that alternatives to the SAP channel are considered, but enough direction must be provided to ensure that the market for DV receiving equipment will not be diluted by multiple incompatible technologies. SAP would be a good common solution in the near term, but stations might preempt DV over the SAP channel for second language broadcasts or other commercial endeavors. VBI would be a more flexible option because it allows a separate audio channel although at a higher cost to the consumer. As people move towards cable TV , it might make sense to offer cable boxes that handle extra audio channels on the VBI. With the advent of ATV and HDTV it will be necessary to ensure that standards for ATV/HDTV, when adopted, incorporate a channel reserved for DV.

COSMOS recommended conducting tests to find out who the DV audience will be. Although their study did not consider the issue of non-visually impaired listeners utilizing DV, that issue may be critical to the commercial viability and subsidy requirements of DV. At least one study should be conducted under a grant or contract to determine DVs commercial viability for both the sensory impaired and non-sensory impaired populations.

6.2 Technologies for Hearing Impairment

6.2.1 Adaptive Modems and TDD Access

Competition in the modem industry will ensure that modem transmission rates double every two to three years, up to the theoretical limits of telephone lines. Data compression will push modems' effective transfer rates well beyond 25,000 bps. Modem manufacturers have the economic incentive to use coding techniques, data compression, and other technologies to push the theoretical limits of data transfer rates. An initiative by the

Department of Education would ensure that these modems will support ASCII and Baudot TDD operations. With research and development action over the next 3-5 years, TDD-capable modems could achieve parity with modems designed for the general public by the mid-1990s. If all new digital signal processing modems sold in the U.S. after 1995 are required to be TDD-capable, businesses and consumers would, for the first time, be able to buy a single-unit ASCII/Baudot modem at competitive prices, with the features and transmission rates available to all modem users. The cost of Baudot and ASCII TDD modems would be amortized over many thousands of units with the benefits of improved communication access for the hearing impaired and greater access to the hearing impaired by the general population. Three to five years after regulatory or legislative action is taken requiring all advanced technology modems to be TDD-capable, several million modems would have been replaced with Baudot TDD-capable ASCII modems and TDD access would be almost universal.

Before TDD compatibility of ASCII modems can be mandated, a standard for TDD modems must be established and technical requirements must be defined. The Department of Education has initiated a TDD modem specification committee through the Lexington Rehabilitation Engineering Center in New York. The Department of Education's next step should be to obtain a draft standard recommendation on TDD services and distribute it to the telecommunications industry, the National Institute of Standards and Technology (NIST) and the FCC.

It is recommended that an Advanced Modem Development Committee be formed with representatives from the telecommunications industry, Department of Education, FCC, and NIST, to determine how the required draft standards can best be implemented. A recommended standard should then be submitted to the FCC for processing within its regulatory charter.

The Department of Education should also consider enacting programs in major cities to ease the transition from TDDs to computers through training programs. Personal computers provide more power for the dollar than TDDs, and personal computers benefit from the continual product improvements and cost advantages that intense industry competition promotes. In general, sensory impaired individuals should be encouraged to use products that serve a large segment of the population as these products become more cost effective than special-purpose devices such as Baudot TDDs.

6.2.2 Telecommunications System Access

The Department of Education should concentrate its efforts on telecommunications system access in the areas of telephone TDD relay services, call progress monitoring, and pay telephone system access. Over the next three to five years the telephone TDD relay systems should mature into a network to serve persons with hearing impairments. The Department of Education should explore methods to expand the use of the network and assist in information dissemination. In addition, the department should explore voice recognition systems to provide input to the TDD relay service to help eliminate operator assistance. This could be done by using speech and natural language systems to translate speech into text and voice synthesizers to translate the TDD text into speech. AT&T begins a speech recognition system in 1992 to provide information services to the hearing populace. In the near future, such systems could provide a relay system to a person equipped with a TDD. All that would then be needed is a speech synthesizer converting the TDD to voice for the hearing person.

The second area is in providing telephone call progress information to persons with hearing impairments. These call progress signals notify persons with hearing when the line is busy, the operator has intercepted the call and when the phone has been off the hook too long, just to name a few. The Department of Education should support device development in this field through the Small Business Innovative Grant Program.

Third, the Department of Education should support research into access to the automatic message answering systems in use by most companies. This includes automatic mail systems as well as menu voice systems used for ordering merchandise or just referring one to a specific service within a company. This research could be done under an innovative research grant.

Development of the technologies required for access to the touch tone signaling and call progress tones would reduce costs to hearing impaired persons in several ways. First, phone calls would not have to be made two or three times to be sure the line is busy or to determine if a TDD is available at the other end. Automated attendant calls, which give prerecorded greetings and voice prompts that render assistance, would be possible since the hearing impaired could use the automated systems for support. This would open up many new sources of information, i.e., customer service, telephone banking, automated ordering systems, etc. The Department of Education should continue to use SBIR and innovative grants to fund research in Telecommunications System Access.

6.2.3 Voice Recognition Systems for Personal and Media Access

The Department of Education should begin the process of developing voice recognition technology for use by persons with hearing impairments by participation in the DARPA-sponsored Speech and Natural Language Workshops beginning in the winter of 1992. This should be followed by the appointment of a Voice Recognition Advisory Committee roundtable to recommend specific goals for developing the technology into devices for persons with hearing impairments for the Department of Education. An extensive applications program should then be initiated to apply the technology to the specific applications defined by the Voice Recognition Advisory Committee, such as closed captioning, TDD relay services, and interpreter services in classrooms and meetings. It is expected that a five year, one million dollars per year effort will be required to develop the technology into prototype products for use by persons with hearing impairments. Innovative Grants, Small Business Innovative Research Grants, and specific applications oriented programs should be initiated to achieve the goals defined by the Department of Education.

The payoff at the end of five years would be to empower the hearing impaired with systems that allow them equal access to television and telecommunications media as well as access to personal communications services.

6.2.4 Video Teleconferencing/Data Compression for Persons with Hearing Impairments

Commercial video teleconferencing, and the use of video compression for all classes of computers, will develop on their own. However, applications that can tolerate lower image fidelity but need extremely high levels of image data compression may develop more slowly.

Specifically, adaptations to provide sign language access through telecommunication systems will tend to develop slowly, so the Department of Education should selectively fund these types of efforts. The Department of Education should continue to support development of edge detection technology specifically for applications to sign language, and should consider joint funding with other Government agencies that are also interested in edge detection.

Likewise, fractal image compression should be investigated, specifically looking for applications for sign language transmission. Fractal image compression development at Iterated Systems began with funding through DARPA, and consultation with them about possible applications to sign language may be helpful. The possibility of joint funding of experiments may also be considered, although it is very important that the special application of sign language be emphasized. Fractal compression itself will develop on its own.

Sign language transmission over computer networks should continue to be explored with funding from the Department of Education.

Commercial video teleconferencing should not receive funding from the Department of Education, because it is profitable in and of itself. However, the Department of

Education should consider working with companies that provide video teleconferencing services to develop a standard sign language terminal for use over phone lines, if those companies show interest in fielding such a product.

6.3 Technology for Visual and/or Hearing Impairments

Potential access improvements with battery technologies should address accessibility to persons with both hearing and vision impairments. Care should be taken in the design of batteries to allow easy installation and replacement by a sensory impaired person. The Department of Education should pay particular attention to types and physical accessibility of the batteries and battery compartments in all designs built under their auspices. It should be easy for a visually impaired person to identify the proper battery in an assistive device, locate and open the battery compartment, and replace the battery with proper orientation. Tactual cues on assistive devices and batteries are necessary to make this possible. Other issues that should be considered include safety, temperature range, reliability , energy density and capacity for recharging.

No battery development by the Department of Education is recommended at this time because other departments of the U.S. Government are pursuing these efforts. The Department of Education should be encouraging better use of battery technology as it is developed to improve performance and/or reduce cost to the sensory impaired population.

APPENDIX A

CONCEPTUAL FRAMEWORK DOCUMENT

EXAMINING ADVANCED TECHNOLOGIES FOR BENEFITS TO PERSONS WITH SENSORY IMPAIRMENTS

CONCEPTUAL FRAMEWORK

March 4, 1991

TABLE OF CONTENTS

Page

1.0 INTRODUCTION 1
2.0 PROGRAM TASKS AND SCHEDULE 1
3.0 THE PROGRAM CONCEPTUAL FRAMEWORK 3
4.0 NEEDS OF PERSONS WITH SENSORY IMPAIRMENTS 7
    4.1 Categories of Impairments to be Considered 8
    4.2 Individual Needs 12
    4.3 Selection Criteria for Advanced Technologies 14
    4.4 Advanced Technology Scenarios 18
5.0 INFORMATION COLLECTION PLAN (TASKS 3, 4, AND 6) . . 19
6.0 PANEL OF EXPERTS/CONCEPTUAL FRAMEWORK (TASKS 2, 5, AND 8) 20
7.0 FINAL REPORT (TASK 10) 21

LIST OF FIGURES

Page

FIGURE 2.0-1. PROGRAM SCHEDULE 2
FIGURE 3.0-1. PROGRAM CONCEPTUAL FRAMEWORK 4
FIGURE 4.3-1. ACTUAL DATA BASE OUTPUT FORMAT 17
FIGURE 7.0-1. FINAL REPORT SCHEDULE 22
FIGURE 7.0-2. FINAL REPORT OUTLINE 23

LIST OF TABLES

TABLE 3.0-1. LIST OF DELIVERABLES 6
TABLE 4.1-1. HEARING IMPAIRMENTS 9
TABLE 4.1-2. HEARING DISCRIMINATION IMPAIRMENTS 9
TABLE 4.1-3. VISUAL IMPAIRMENTS 9
TABLE 4.3-1. SAIC CORPORATE TECHNOLOGY DIVERSITY . . . 16
TABLE 4.3-2. SAIC'S DATA BASE ACCESSIBILITY 17
TABLE 4.4-1. ADVANCED TECHNOLOGIES 19

CONCEPTUAL FRAMEWORK

1.0 INTRODUCTION

The initial step in establishing an advanced technology program was to name the Department of Education's Office of Special Education Programs (OSEP) to define the program goals and focus the effort on specific sensory impairments. The second step was to initiate a program to look at "Advanced Technologies for Benefits to Persons with Sensory Impairments." The logical third step was to select a diversified advanced technology company with the experts, skills and resources to examine advanced technologies in both the private and public (U.S. Government Laboratories) sectors. Science Applications International Corporation (SAIC) is a diversified advanced technology company with the understanding of the technologies being applied within industry, research engineering laboratories, the military, and academia to perform this Department of Education study. In this conceptual framework, SAIC's technical staff has structured a comprehensive program in such a way that it guides future project activities and leads to the development of 10 - 20 advanced technology scenarios. These scenarios will explain how advanced technologies can be used to benefit persons with sensory impairments by improving existing devices and influencing future designs.

2.0 PROGRAM TASKS AND SCHEDULE

The program Schedule is shown in Figure 2.0-1. SAIC's first task was to attend a meeting with the Contracting Officer's Technical Representative (COTR) and the Education Department's Contracting Office in Washington, D.C., within 10 days of the contract award date. SAIC reviewed the goals and objectives of the procurement, SAIC's approach to the major tasks in the procedural plan, and the potential outcomes. SAIC's Principal Investigator, Mr. Daniel E. Hinton, Sr., and the Conference Center's Principal Investigator, Dr. Carl Jensema, provided the Department of Education Contracting Officer's Technical Representative (COTR), Mr. Ernie Hairston, with specific performance goals, a proposed timeline, and an understanding of the Conference Center's and SAIC's corporate resources to accomplish the work.

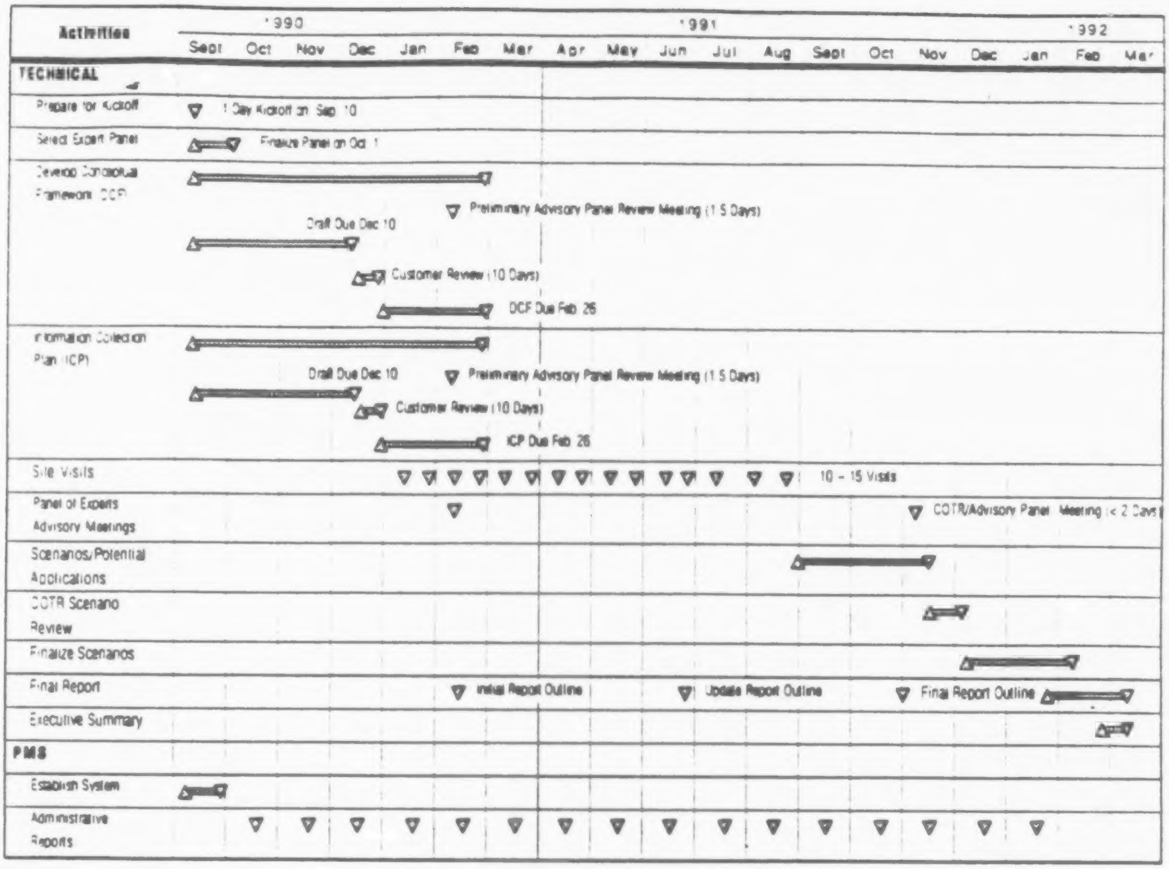


FIGURE 2.0-1. PROGRAM SCHEDULE

Next, the members of the program's Panel of Experts were selected based on their expertise in relevant technologies, vocational rehabilitation or consumer products. The COTR was notified by letter, 15 days after contract award, of the composition of the Panel of Experts. Finally, this "Examining Advanced Technologies for Benefits to Persons with Sensory Impairments" Conceptual Framework and Information Collection Plan (ICP) were developed for the COTR's and the Panel of Experts' review and comment by the third month of the contract. Final discussion and refinement will take place at the Panel of Experts meeting in the fourth month of the contract.

Finally, the Site Visit, Performance Measurement, and Administrative Reports delivery dates were established and the schedule in Figure 2.0-1 constructed using the contract award date of September 10, 1990, as the start date.

BEST COPY AVAILABLE

3.0 THE PROGRAM CONCEPTUAL FRAMEWORK

SAIC's Conceptual Framework integrates the program based on our original proposal submitted to the Department of Education on June 22, 1990. This Conceptual Framework integrates ideas, techniques, technologies, and system concepts from many diverse sources into a program to meet the needs of persons with sensory impairments. SAIC recognizes that the diversity and sheer number of alternatives that could be generated during this effort places a burden on program management and increases the risk of the effort. The study process as presented in the Statement of Work (SOW) represents a significant risk reduction measure and is, therefore, implemented in this Conceptual Framework. SAIC is pursuing a disciplined approach to organizing and implementing the study as outlined in this Conceptual Framework.

SAIC's methodology for program task execution is presented in the Program Conceptual Framework, Figure 3.0-1. The Tasks consist of three types:

- Program management and control,
- Expert advice and oversight, and
- Study execution tasks.

Program management and control tasks include program planning (Task 1 - COTR meeting within 10 days of contract award) and study performance measurement (Task 11 - Establish A Performance Measurement System). Our approach is to use this conceptual framework as a program plan and use management controls, both routine SAIC management techniques and COTR-approved performance measurement approaches, to manage program implementation. It is particularly important to manage risk in the planning tasks (Task 3, Develop A Conceptual Framework, and Task 4, Draft An Information Collection Plan) and in the selection of data from which to develop a limited number of high-quality scenarios. The Performance Measurements System was implemented at program award and briefed to the COTR at the COTR meeting 10 days after contract award. This document is the Conceptual Framework with the Information Collection Plan as Attachment 1.

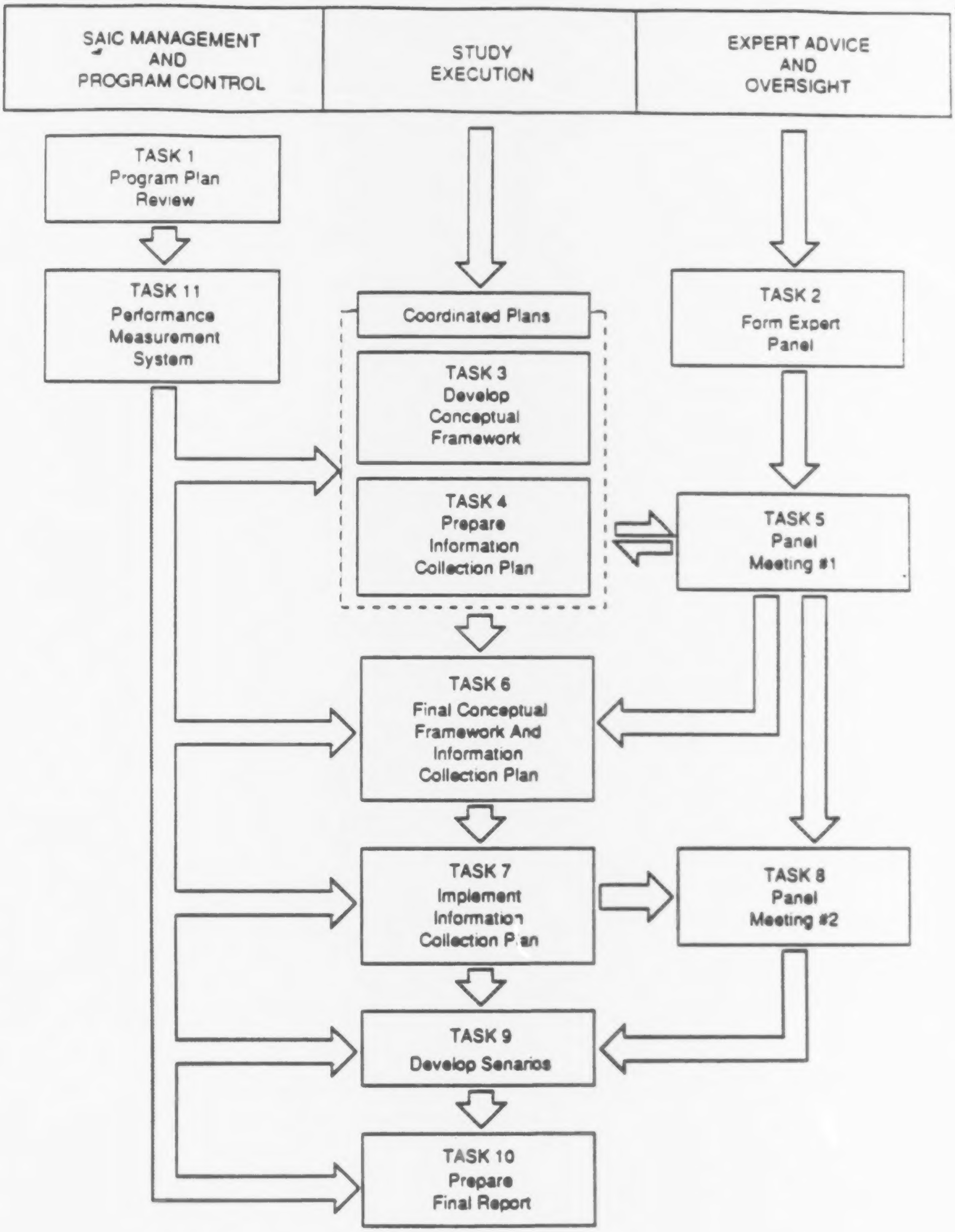

SAIC MANAGEMENT AND PROGRAM CONTROL
STUDY EXECUTION
EXPERT ADVICE AND OVERSIGHT
TASK 1 Program Plan Review
TASK 11 Performance Measurement System
Coordinated Plans
TASK 2 Form Expert Panel
TASK 3 Develop Conceptual Framework
TASK 4 Prepare Information Collection Plan
TASK 5 Panel Meeting #1
TASK 6 Final Conceptual Framework And Information Collection Plan
TASK 7 Implement Information Collection Plan
TASK 8 Panel Meeting #2
TASK 9 Develop Senarios
TASK 10 Prepare Final Report

FIGURE 3.0-1. PROGRAM CONCEPTUAL FRAMEWORK

Expert advice and oversight tasks are centered on the Panel of Experts. SAIC will use expert advice to guide the technical efforts and provide feedback to management control as a way of controlling risk, measuring progress, and planning expenditures. The tasks from the statement of work are as follows:

- Task 2 Form Panel of Experts.
- Task 5 Schedule a 1 1/2-day meeting with COTR and Panel of Experts convening in Washington D.C. area in fourth month of the contract.
- Task 8 Schedule a 2-day meeting with COTR and Panel of Experts convening in Washington D.C. area in 14th month of the contract.

The following study execution tasks form the core of the study to be performed by the SAIC staff:

- Task 3 Develop Conceptual Framework to Guide Project Activities by the third month of the contract for review by the Panel of Experts and the COTR.
- Task 6 Submit to the COTR the final Conceptual Framework and Information Collection Procedures in the fifth month of the contract.
- Task 7 Implement the Procedures outlined in the Information Collection Plan with a minimum of 10 and a maximum of 15 site visits.
- Task 9 Develop up to 20 Scenarios featuring potential applications of existing technology and aspects of technology that show promise for facilitating the access of individuals with sensory impairments to media and communications.
- Task 10 Prepare a Final Report and Ten Year Development Plan.

The implementation methodology is to draw on inputs from management and expert advice task outputs. The planning and implementation tasks are reviewed and guided by the advice of the Panel of Experts and the Department of Education's COTR. The relationship between the Panel of Experts and SAIC's technical Staff is shown in Figure 3.0-1. The arrows indicate an interactive relationship with a logical flow of advice and information between the SAIC team and the Panel of Experts. As the SAIC staff formulates the Information Collection Plan and identifies technologies for review at the Panel of Experts meetings, the Principal Investigator will call specific panel members for

TABLE 3.0-1. LIST OF DELIVERABLES

ITEM	DATE DUE	QUANTITY
List of Advisory Board Members (Task 2)	October 1, 1990	2
Administrative Reports (Task 11)	Monthly. See Figure 2.0-1	2
Draft of Conceptual Framework (Task 3)	December 10, 1990	2
Draft of Information Collection Plan (Task 4)	December 10, 1990	2
Final of Conceptual Framework and Information Collection Plan (Task 6)	February 26, 1991	2
List of Organizations Technologies (Task 6)	February 26, 1991	2
Case Report from Site Visits (when applicable) (Task 7)	September 10, 1991	2
Case Report from site visits (when applicable) (Task 7)	October 25, 1991	2
Case Report from site visits (when applicable) (Task 7)	December 10, 1991	2
Draft of Scenarios (Task 9)	December 10, 1991	2
Draft Final Report (Task 10)	February 10, 1992	2
Final Scenarios (Task 9)	February 10, 1992	2
Final Report and Ten-Year Development Plan (Task 10)	March 10, 1992	2

advice on technologies and their applicability to specific sensory impairments. Program control exercises oversight as the time-phased tasks are executed, leading to the final report of findings for the study. The COTR oversight was established through the formal contract. This relationship is based on the deliverables in Table 3.0-1.

The deliverables provide continuous program monitoring and control by the Department of Education's Office of Special Education Programs (OSEP) COTR. In addition, copies of the deliverables provide the Panel of Experts with a basis for expert advice and oversight.

4.0 NEEDS OF PERSONS WITH SENSORY IMPAIRMENTS

To begin implementing the Program's Conceptual Framework one must first understand the basic needs of persons with sensory impairments. These cannot be the preconceived needs of the technical staff or academic experts, but, rather, the needs as expressed by persons with sensory impairments. SAIC's Principal Investigator, Mr. Hinton, has worked closely with persons with sensory impairments over the past ten years. From this work, a set of needs have been expressed by persons with sensory impairments. Persons with sensory impairments have the same basic needs as the population as a whole, such as a comfortable place to live, meaningful employment, and opportunities for recreation and socialization based upon life style options and individual choice. To meet these needs, persons with sensory impairments require aids and devices that expand their access to media, information and communications capabilities. This access has the potential for expanding the individual's options and choices in vocation, recreation, and lifestyle. For example, a real-time court room stenographic speech-to-text translation system might allow persons with hearing impairment to practice law in court. In general, the aids and device needs of persons with sensory impairments can be grouped into two areas:

1. Physical use (i.e., communications, mobility, and situational awareness), and

2. Personal use (i.e., vocation, education, recreation, and life style).

These aids and device groupings overlap since the physical use devices can serve the same purpose as the personal use devices in providing equal access for persons with sensory impairments. The internal or external operation of the advanced technology information and communications devices needed to meet these needs can be extremely complicated, but the devices' human factors (i.e., the interface presented to persons with sensory impairments) must be functional and simple to understand and operate. An example of a technology that meets these requirements is the video camera being sold to the general public, adapted for large character displays for persons with vision impairments. Although the technology is complicated--using microelectronics, advanced optics, charge-coupled devices, and advanced materials technologies--the operator only needs to point the camera and press the trigger to record an image or view the result on a large character television monitor. Although the underlying technologies that are used cost several hundred million dollars to develop, the economies of scale in producing one million or more cameras make the final product cost less than $600 to $1,000 each at the retail level. In the past, specially developed camera systems for persons with vision impairments cost $5,000 to $10,000.

4.1 Categories of Impairments to be Considered

To enumerate all the different categories or combinations of sensory impairments is not the intent of this Conceptual Framework. Tables 4.1-1, 4.1-2, and 4.1-3 provide a technical overview of hearing loss, hearing discrimination, and vision sensory impairments, respectively. However, an understanding of these tables, in relationship to advanced technology and how it can be applied to persons with sensory impairments, is essential to program implementation.

Tables 4.1-1 and 4.1-2 illustrate that hearing impairments for this study must be considered in the context of the adaptive needs of the individuals with hearing impairments. If the only problem was sensitivity to sound volume, then hearing aids or body aids would be an adequate solution. There would be little need for this study to examine applications for persons with hearing impairments. However, discrimination ability and profound sensitivity loss must also be considered by engineers and scientists examining technologies

TABLE 4.1-1. HEARING IMPAIRMENTS

Decibel Loss	Hearing Loss	Typical Sound
16-20	Slight	Whisper
26-40	Mild Soft	Speech
41-55	Moderate	Loud Speech
56-70	Moderately Severe	Loud Music
71-90	Severe	City Traffic
91 or more	Profound	Loud Rock Band/Chain Saw

Note: This table is a measure of sensitivity to sound and not the ability to discriminate or understand speech. If there is only a loss of sensitivity, then a hearing aid can help.

TABLE 4.1-2. HEARING DISCRIMINATION IMPAIRMENTS

Discrimination Ability	Hearing Problem
75-90%	Mild difficulty understanding speech
60-75%	Moderate difficulty in communications
5-60%	Moderately severe difficulty in communication
Below 5%	Severe difficulty in communication

Note: Simply amplifying sound does not necessarily help a discrimination problem, but, when there is a drop in sensitivity, a hearing aid can be beneficial.

TABLE 4.1-3. VISUAL IMPAIRMENTS

Measures	Visual Problem
20/20 (Snellen Score)	Perfect Vision
20/200 With Best Correction	Legally Blind
20 Degrees in Eye With Best Vision	Visual Field Legally Blind
Special Condition	Color/Night/Snow Blindness

in addition to sound level sensitivity. The multi-faceted aspects of hearing require applying advanced technologies in many fields of science and engineering, from microelectronics to optical displays.

For those individuals with hearing sensitivity loss, advanced technologies, such as: microelectronics; acoustic microphones; and materials that allow the aid to survive in the ear channel, will be examined in this study. In particular, military specifications and implementations for severe environments are applicable. Other technologies, such as digital signal processing for background noise suppression, and signal processing based on the monaural and binaural frequency loss of individuals, will also be examined.

For those individuals with profound hearing loss or discrimination ability loss, this study will examine the following technologies:

- Automatic speech recognition (ASR)
- Modem technologies for telecommunications access
- Cellular telephone technologies for individual communications devices
- Close caption technologies as applied to video systems

Also, sound processing technologies that can be adapted to an individual's particular impairment will be explored. This will require an examination of software and hardware technologies for sound processing and amplification.

This study will include, but not be limited to the capability to process natural voice. This includes the processing of voice patterns of persons with a hearing impairment to allow others to better understand them, both in person and on the phone (i.e., automatic voice amplification and word reconstruction).

The problems of persons with vision impairments shown in Table 4.1-3, vary from those with total blindness to those with limited vision or specialized vision losses (e.g., night, color, etc.). The technologies that will be addressed to assist persons with vision impairments are:

- Special optics (e.g., glasses, magnifying devices, and large print)
- Text page scanners
- New optical systems (e.g., "Private Eye" that is only 3 inches square and projects a 10 inch picture in front of the user)
- Character recognition software and hardware
- Vision enhancement (e.g., infrared devices, color recognition devices)
- Special materials to allow small, inexpensive braille devices (e.g., superconductor magnets for braille print heads and electronic page braillers)

SAIC will address manufacturing technologies in each technology scenario that is developed because devices must be produced at a cost that persons with sensory impairments can afford. Based on his detailed knowledge of manufacturing technologies, Dr. Kelly, Department of Defense Advanced Research Projects Agency (DARPA), was invited to serve on the Panel of Experts, to provide guidance on addressing manufacturing technologies. Estimates of manufacturing technology cost, manufacturing process cost and wholesale and retail costs will be included. SAIC understands that any technology that costs more to produce than persons with sensory impairments can afford will not be applied to meet a need, unless there is government or private assistance available to offset the cost. To justify the cost, the device must provide a job-related benefit to empower the person to become self-sufficient.

The general impairment category that has been most publicized with respect to advanced technology is physical impairments, where advanced technology has been applied to improving wheel chairs, adapting automobiles, and public transportation systems for access. Advanced technology has been most successful when the devices have been designed to allow access by a broad range of intellectual levels, from persons with average intelligence to persons with various learning disorders. The key is simplicity of design and function. For example, a lightweight carbon-graphite wheelchair with a motorized control system allows paraplegics to have physical access through increased mobility. However, the chair's controls are a communications device that allows the user access to the system using hand, voice, eye or muscle movement.

Information and communications media access aids, though important, do not receive the attention they deserve. This is because they are not highly visible to the general public since they are personal use items, restricted to home, office or academic environments. Close caption sets, for example, allow persons with hearing impairments access to television and communication media by employing advanced microelectronics and video technologies. The general public is seldom exposed to the technology or the benefits experienced by persons with hearing impairments because the technology is transparent to the general population. Since the general public and equipment manufacturers are not aware of the need for media access for persons with sensory impairments, media access is not being addressed by the companies and government organizations responsible for developing the next generation of devices. An example is High Definition Television (HDTV) standards, television studio equipment, and home television systems. Without awareness of the potential problem, the general public or persons with sensory impairments may not petition for government action to require close captioning to be a part of every HDTV set through the adoption of a standard providing for equal HDTV media access.

4.2 Individual Needs

Persons with sensory impairments have special needs that require individual solutions based on the degree of sensory impairment (i.e., degree of hearing or vision loss, time of onset of the loss, physical and learning ability). The problem with individual solutions for the problems associated with vision and hearing loss is the cost of developing the required technology (i.e., cost of development vs. number of persons served). To amortize the cost over a larger population and thus justify the cost, research and development in the past has been directed towards meeting the needs of persons with sensory impairments that could be grouped into a large population. In many cases, the objective has been general research and not technological applications. One notable exception is closed captioning for the hearing impaired as discussed above. The development of video technology to allow the hearing impaired to read television has met the needs of over 2 million people. Again, the goal was to meet the needs of the largest hearing-impaired population. Those with the dual impairments of hearing impairment and low vision or selective vision (i.e., color vision disorders) were excluded because the decoders developed could not adjust the character size or color to meet the needs of a dual disability.

This advanced technology program's goal is to identify practical ways to meet the needs of persons with single and multiple impairments through technology application scenarios. The recommended technologies will allow adaptation through software or plug-in modular hardware to make use of the equipment designed for a larger population. This study will address multiple solutions to meet individual needs of persons with sensory impairments.

Computer technology is a good example of advanced technology that has been adapted to meet the media access needs of persons with sensory impairments. For the hearing-impaired, information and communications exchange is now possible, between individuals or with computer databases, using standard modems. For the visually impaired, speech modules take the place of a screen by allowing the words or characters stored in screen memory to be spoken. Although these advanced technologies are inexpensive, they do not meet all the information and communications needs of persons with sensory impairments. Computer modems for the general public, costing $50 to $100, do not allow communications with the TDD devices for persons with hearing impairments. Special modems have been developed, costing $250 and more, that allow TDD and standard modem operation. However, these special modems still limit information and communication exchange rates with other computers to 1,200 to 2,400 bits per second while in the standard modem mode.

4.3 Selection Criteria for Advanced Technologies

The selection criteria used for selecting the technologies to be examined in this study were near term (3-5 years) and far term (5-10 years) impacts on the information and communication needs of persons with sensory impairments. The SAIC team, with the assistance of the Panel of Experts and guidance from the Department of Education COTR, will apply four criteria in determining the final list of technologies and scenarios to be examined as follows:

- Does the technology apply to a specific information or communication need of persons with hearing or vision impairments? (e.g., telephone communication, newspaper text scanner, etc.)

- Can the technology be applied to the specific information or communications needs within the next three to five years? (e.g., HDTV closed caption broadcasting, voice recognition, voice reprocessing, braille cell production, etc.)

- Can the technology be applied to the information or communication needs within the next 10 years? (e.g., voice recognition without breaks between words)

- Is it feasible to apply the technology to the information or communication needs of persons with sensory impairments? (i.e., cost, size, weight, power, etc.)

The criteria used to select fifteen to twenty organizations from which to collect information on the technologies is as follows:

- Does the organization have a long-term commitment to the technology? (i.e., past developments, being organized to research, develop, and produce the technology, etc.)

- Does the organization have the technical staff to support the present and future developments? (i.e., engineers, technical staff, etc.)

- Is there a stated commitment to support the technology at a high level? (i.e., management commitment)

The concept is to interest each organization in this study. SAIC started the process by forming a distinguished Panel of Experts representing each area from consumer groups, industry, government, academia and rehabilitative engineering centers throughout the United States. This will facilitate the study, since experts from several of the target organizations with interest in the advanced technologies are on the Panel of Experts. The Panel of Experts members will be requested to make personal introductions within a company, industry or government agency where a specific technology is being investigated and they have a professional contact. SAIC's technical staff's extensive contacts will also be applied.

SAIC has diverse corporate resources located throughout the United States and these resources will be brought to bear on each technology field. Specifically, SAIC has divisions that are devoted to most of the military, energy and business technologies that will be explored in the study. SAIC's corporate diversity is shown in Table 4.3-1. Company experts will be used to help develop two-to-five page advanced technology scenario papers on specific technologies and ways to apply these technologies to the information and communications needs of persons with sensory impairments. These experts will help identify specific individuals and companies and make formal contacts with the industrial experts working on the next generation technologies. These papers will first be developed into scenarios and then into a short synopsis paper for dissemination by the Department of Education.

A key element of our program is a database search on the advanced technologies. SAIC's Corporate Technical Resource Acquisition Center (CTRAC), located in McLean, Virginia, will conduct literature searches on key words related to the technologies applicable to information and communications media access needs of persons with sensory impairments based on the technologies and keywords identified by SAIC's technical staff. Table 4.3-2 is a partial list of databases that can be searched with on-line computer services. Figure 4.3-3 illustrates the format SAIC will use to create our database from the on-line and off-line searches. SAIC already has on hand several hundred companies

TABLE 4.3-1. SAIC CORPORATE TECHNOLOGY DIVERSITY

Artificial Intelligence	Sensor Technology	Safety Technology
Optical Data Processing	Acoustics	Telecommunications
Man-Machine Interfaces	Automation	Aerospace Materials
Magnetics	Interactive Computer Systems	Optical Sciences
Cybernetics	Fiber Optics	Signal Analysis
Semiconductor Technology	Human Factors	Computer Displays
Electro-Optics	Signal Processing	Digital Data Transmission
Behavioral Sciences	Computers	Aerospace Structures
Materials Sciences	Space Technology	Nondestructive Testing
Applied Mechanics	Materials Testing	Fluid Mechanics
Fluid Physics	Continuum Mechanics	Theoretical Chemistry
Plasma Physics	Theoretical Physics	Information Theory
Detection Phenomena/Systems	Gas Dynamics	Power Engineering
Information Processing	CAD/CAM	X-Ray Diagnostics
Thermodynamics	Process Instrumentation	Solid-State Lasers
Electromagnetic Propagation	Laser Diagnostics	Laser Beam Propagation
Pointing and Tracking Systems	Beam Physics	Plasmascope Displays
Low-Light-Level Detectors	Life Sciences	Industrial Hygiene
Environmental Health	Carcinogenic Substances	Bioluminescence
Toxicology	Robotics	Diagnostics
Radiography	Elemental Analysis	Chemiluminescence

TABLE 4.3-2. SAIC'S DATA BASE ACCESSIBILITY

On-Line:	Newspaper/Law (Nexis/Lexis) The Scientific and Technical Newwork (STN) On-Line Computer Library Center (OCLC) Remote Console = NASA On-Line (Recon) MEDLARS (National Library of Medicine's Medline, Chemline, and Toxline) Defense Research On-Line (RECON) United States Naval Institute (USNI) NTIS/DTIC (Classified Documents Databases Search Capability)
Off-Line:	The Gold Book (Guide to Manufacturers and Technologies) The Sensors Buyer's Guide (Indexed by Manufacturer, Property, Technologies)

Company:	Digital Design
Address:	Industrial Vision Division 3060 Business Park Drive Norcross,Georgia 30071
Phone #:	404/447-0274
FAX#:	404/263-0405
Established:	1981
Contact:	Francois Perchais, Manager Industrial Vision
Properties:	Vision/Image Sensing, Light (IR and Visible)
Technologies:	Charge Coupled Devices (CCD); Lasers; Optical; Optoelectronic; Phototransistor/Diode
Related Products/Services:	Vision Systems; Computer Software for Interfacing and Applying Sensors; Custom Design; Data Acquisition; Signal Processing
Site Visit:	(List Date and Time for Site Visit)

FIGURE 4.3-1. ACTUAL DATA BASE OUTPUT FORMAT

cataloged in this format but will do a database search on the scenario applications. CTRAC's analysts will assist the SAIC Principal Investigator and the program team to rapidly identify key organizations to visit and gather information for developing the scenarios necessary to meet the goals and objectives of this critical Department of Education program. In addition to database searches, CTRAC, as a member of the Special Libraries Association, American Society For Information Sciences, and Interlibrary Users Association can obtain almost any article or book written in the world, including translations of foreign documents.

4.4 Advanced Technology Scenarios

The definition of advanced technology is arbitrary since many technologies must be combined to meet the needs of persons with sensory impairments. The traditional advanced technology areas are optics, microelectronics, materials (i.e., graphite composites and metals), and biomedical. Table 4.4-1 provides a representative list of those technologies that have a direct or integrated application for persons with sensory impairments. During the course of the program, SAIC will explore these advanced or emerging technologies and develop scenarios for applications which benefit persons with sensory impairments. Specifically, access to communications media such as films, video, television, print, telecommunication devices, electronic correspondence, innovative uses of current communications devices (facsimile, computers, page scanners, etc.) will be considered. In addition, SAIC's technical staff will develop the scenarios to identify specific applications and features of applications that facilitate or limit media communications access of individuals with specific disabilities. This effort will include but not be limited to identifying:

- adaptations that facilitate access and minimize barriers,
- the development and evaluation activities necessary to achieve those adaptations, and
- the number and type of groups benefitting from the technology applications.

5.0 INFORMATION COLLECTION PLAN (TASKS 3, 4, AND 6)

The Information Collection Plan is Attachment 1. This plan fulfills the Statement of Work tasks as follows:

- Task 3 By the third month of the program develop a Conceptual Framework to guide project activities. This Conceptual Framework will be reviewed by the Panel of Experts and the COTR.

TABLE 4.4-1. ADVANCED TECHNOLOGIES

High Definition Television	Closed caption Audio captioning Image enhancement and enlargement
Automatic Speech Recognition (ASR)	Continuous recognition Environmental noise reduction
Processing of Natural Voice	Adjust speech for recognition by hearing community Volume control for telephone conversations Signal processing to filter out background noise
Optical Automatic Character Recognition	Page scanners Vision aids (street sign reader) Image enhancement (NASA's Low Vision Enhancement System) Display recognition
Neural Network Technology	Automatic speech recognition Voice recognition Automatic character recognition
Telecommunications	Cellular radio combined with modems (TDD), natural voice recognition from central computer or personal computer modems, digital signal processors and software to be compatible with TDD. Interactive video Very small satellite systems Fiber optics impact Data compression impact
Microelectronics	Miniaturization of aids and devices New devices (e.g., high-temperature superconductors
Amplification System	Noise reduction Adaptive aids

- Task 4 Draft an Information Collection Plan.

The Information Collection Plan includes:

  - A list of 15 organizations for site visits, including rehabilitation engineering centers, private industry, the Department of Defense (DoD), and the National Aeronautics and Space Administration
  - A list of technologies and aspects of technology in general, to be investigated within or across these organizations

- Task 6 In the fifth month of the contract, submit the final Conceptual Framework and Information Collection Procedures to the COTR for approval.

6.0 PANEL OF EXPERTS/CONCEPTUAL FRAMEWORK (TASKS 2, 5, AND 8)

A key feature of this project is the panel of nationally known experts in areas related to technology and persons with sensory impairments. The purpose of the Panel of Experts is to bring together professionals and persons with sensory impairments representing not only great depth of technical knowledge, but also extensive understanding of the needs of persons with sensory impairments. Through open discussion with the members of this panel, the project staff expects to obtain invaluable guidance in their efforts to pinpoint relevant technological areas and scenarios for examination. It is the advisory panel's responsibility to provide guidance and critical evaluation of the project staff's plans, research activities, and interpretation of data.

The advisory panel for this project is organized and managed by the Conference Center, Incorporated, which specializes in research, training, and conference management involving issues related to hearing and visual impairment. The Conference Center's work in this project is headed by Dr. Carl Jensema, a nationally recognized authority on technology for persons with a variety of sensory impairments.

Although the focus of the Conference Center's work will be the advisory panel, Dr. Jensema has been, and will continue to be, personally involved in all phases of the project.

7.0 FINAL REPORT (TASK 10)

SAIC's technical staff views the final report on "Advanced Technologies for Benefits to Persons with Sensory Impairments" as the beginning of the Department of Education's future advanced technology program. The SAIC final report will provide a baseline for exploring advanced technologies to meet the information and communications needs of persons with sensory impairments.

The report will minimize technical jargon and focus on applying technology to media access. Thus, it will provide a comprehensive understanding of the program methodology and execution and serve as a road map to assist the Department of Education in assessing the use of advanced technologies for the sensory impaired.

The final report schedule is shown in Figure 7.0-1. The initial outline of the final report shown in Figure 7.0-2, will be prepared by the fifth month of the contract and provided in the administrative report to the COTR. The outline will be updated in the 11th month of the contract. Finally, a draft of the final report will be provided to the COTR for review in the 17th month of the contract. The final report will then be submitted in the 18th month of the contract in fulfillment of the contract.

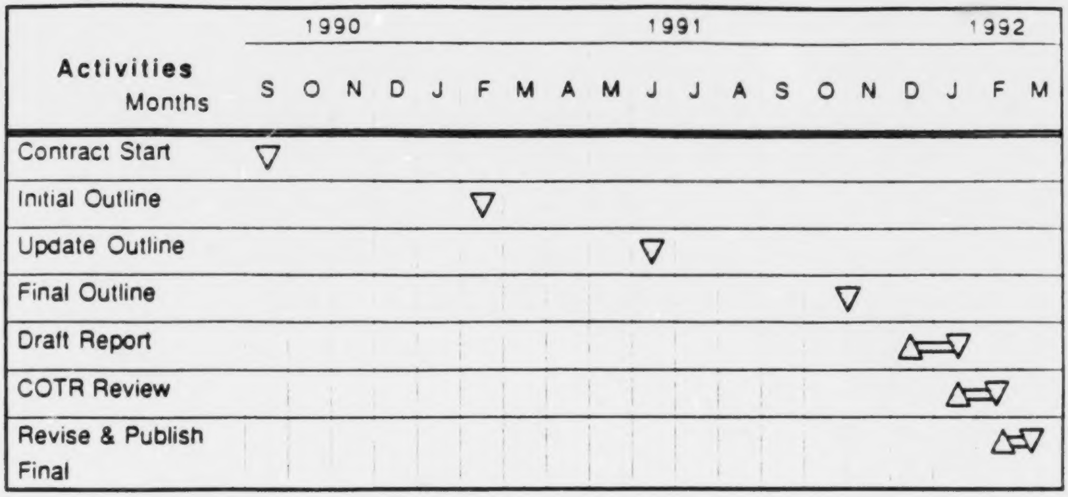


FIGURE 7.0-1. FINAL REPORT SCHEDULE

The final report will include the relevant issues in advanced technologies, the applications, their estimated costs, and other factors that would enhance or limit potential advanced technology applications. An appendix to the report will describe a 10-year action plan designed to help the Department of Education to emphasize the most urgent access requirements.

EXECUTIVE SUMMARY

This will be a 10-page summary intended for executive-level decision makers who need specific information on advanced technology trends that offer the most benefit for persons with sensory impairments. It will project potential outcomes over the next 3-5 years, and 5-10 years, with encouragement from the Department of Education. The principal topics for the study will be covered, including a clear, concise statement of the approach used, and a discussion of the findings and specific recommendations on advanced technologies to be addressed for various target audiences. A matrix of the scenarios versus technologies will also be included that shows at a glance the potential outcomes and benefits for persons with sensory impairments.

1.0 **Introduction**

This will establish the framework of the project for the final report. The introduction will include the structure of the report, the concept of the study and the most significant outcomes.

2.0 **Purpose and Objective of the Project**

SAIC will clearly state the purpose and objective of the project.

3.0 **Approach Employed**

The approach to program execution will include:

- Organization and conduct of the Panel of Experts meetings
- The approach to database searches
- The approach to data collection and site visits
- The approach to scenario development

4.0 **Results and Findings**

This will be a discussion of the results and findings of the project as they relate to advanced technologies for benefits to persons with sensory impairments. The results and findings will address specific needs and applications to meet the needs of persons with sensory impairments.

5.0 **Conclusions**

The conclusion will address the scenarios and how they relate to persons with sensory impairments. An estimate of the value of the scenarios to the target audience will be projected for 3- to 5- year and 5- to 10- year timeframes.

6.0 **Recommendations**

Specific recommendations will be made about the need for Department of Education involvement or sponsorship of a particular advanced technology for benefits to persons with sensory impairments.

Appendix A: Ten-Year Development Plan

SAIC will provide a ten-year advanced technology action plan to assist the Department of Education in developing priorities that meet the most urgent media access needs of the sensory impaired.

FIGURE 7.0-2. FINAL REPORT OUTLINE

APPENDIX B

INFORMATION COLLECTION PLAN

EXAMINING ADVANCED TECHNOLOGIES FOR BENEFITS TO PERSONS WITH SENSORY IMPAIRMENTS

INFORMATION COLLECTION PLAN

March 4, 1991

Table of Contents

1.0 INTRODUCTION 1

2.0 BACKGROUND 2

3.0 INFORMATION COLLECTION PLAN CONCEPTUAL FRAMEWORK . 2

4.0 SITE VISITS 9

5.0 PROCEDURES TO GAIN SITE ACCESS 13

6.0 INFORMATION COLLECTION PROCEDURES 14

7.0 OVERCOMING POTENTIAL BARRIERS 15

LIST OF FIGURES

1.0-1. Information Collection Plan Conceptual Framework 1
4.0-1. Federal Laboratory Consortium Technical Request Form 10
4.0-2. Sample Scenario Outline 11
4.0-3. Form for Information Collection Worksheet/Questionnaire 12
7.0-1. Example Letter to Company on Overcoming Proprietary Barriers . 17

LIST OF TABLES

3.0-1. Suggested Technology Scenarios 3
3.0-2. Databases Searched by SAIC 4
3.0-3. Journals SAIC has Searched 5
3.0-4. Trade Shows SAIC Has Attended or Is Scheduled to Attend 6
3.0-5. Contacts Made at the National Home Health Care Expo 7
3.0-6. Contacts Made at the NASA Tech 2000 Show 8

LIST OF TABLES: ENCLOSURE 1

1A. Companies/Researchers to Contact 19
1B. NASA Technology Utilization Officers to Contact 26
1C. Federal Laboratory Consortium Contacts 28
1D. Defense Department Laboratory Contacts 28
1E. Department of Education Engineering Center Contacts 29

INFORMATION COLLECTION PLAN

1.0 INTRODUCTION

This Information Collection Plan (ICP) outlines SAIC's specific approach to collecting information on advanced technologies, ranging from unstructured telephone interviews to structured site visits.

An Information Collection Plan Conceptual Framework is shown in Figure 1.0-1. The Information Collection Plan includes the following elements:

- A delineation of how sites were selected for the study
- Procedures being used to gain access to the sites
- An outline of what information will be collected at each site
- How the information will be collected
- How potential barriers to information collection are being overcome

Figure 1.0-1. Information Collection Plan Conceptual Framework

The Information Collection Plan is divided into 7 sections. Section 1 outlines the document that follows. Section 2 sketches the experience that SAIC is able to bring into this study. Section 3 specifies the databases, journals, and other sources SAIC has consulted, and trade shows attended. These resources serve both as sources of information on relevant technologies and to determine what sites to contact and/or visit. Section 4 describes the sites to be visited and preparation for those site visits. Section 5 outlines procedures to gain access to those sites, and section 6 explains the results of site visits. Finally, section 7 describes how SAIC is overcoming proprietary and classification barriers to information collection.

2.0 BACKGROUND

The SAIC Information Collection Plan is based on our corporate involvement with key advanced technology industries, government organizations, rehabilitation engineering centers, and universities. As discussed in the Conceptual Framework, SAIC's Panel of Experts includes representatives from industry, defense, rehabilitation, and the academic community. When necessary, SAIC's Principal Investigator will request members of the Panel of Experts to recommend sources and contacts for information collection.

SAIC's Principal Investigator, Mr. Hinton, has been identifying and cataloging key technologies that have applications to persons with sensory impairments for the past ten years. His ongoing efforts to identify technologies have been directed toward providing aids to allow persons with deaf-blindness to gain access to information and communication systems. Technologies that have been applied include infrared and optical technologies, the Deaf-Blind Computer Terminal Interface (computer and braille technologies), the voice modulation device indicator (microelectronics technology), and the Braille Telecaption System (video, computer, and braille technologies). Mr. Hinton's extensive personal files are being used to narrow the field to specific high technologies to be examined for this program.

3.0 INFORMATION COLLECTION PLAN CONCEPTUAL FRAMEWORK

In the ICP Conceptual Framework, shown in Figure 1.0-1, the first step was to identify specific needs of persons with the sensory impairments in the program's Conceptual Framework and the relevant technologies to meet these needs. The candidate technology scenarios for consideration by the Panel of Experts are shown in Table 3.0-1.

Table 3.0-1. Suggested Technology Scenarios

Technologies for Visual Impairments	
1	Braille Devices and Techniques to Allow Media Access
2	Input/Output Devices for Computer & Electronic Book Access
3	Visible Light Spectrum Manipulation to Allow Media Access for Persons with Selective Vision
4	Flat Panel Terminal Displays Used with Page Scanners
5	Sign Access (Talking Signs & Video Techniques for Reading Printed Signs)
6*	Character Readers for Dynamic LED and LCD Display Access
7*	Descriptive Video for Television Access
Technologies for Hearing Impairments	
1	Adaptive Modems and TDD Access
2	Telecommunications System Access (Touch Tone Signaling Access)
3	Voice Recognition Systems for Personal and Media Access
4	Video Teleconferencing/Data Compression for Persons w/Hearing Impairments
5*	Continuous Speech Recognition for Real-Time Closed Captioning (of Television and Video Media)
Technologies for Visual and/or Hearing Impairments	
1	Portable Power Systems
2*	Emergency Warning Devices for Emergency Systems Media Access
3*	Cellular Telephone Media Access
Low Priority Technologies for Visual and/or Hearing Impairments	
1*	Natural Voice Processing for Telephone Access
2*	Input and Output Devices for Reading and Displaying Sign Language
3*	Voice Synthesis Systems for Media Access

* To be examined if time and resources permit.

The ICP Conceptual Framework, shown in Figure 1.0-1, will be reviewed by the Panel of Experts, and the Panel will provide recommendations on the candidate technology scenarios and sites to be visited from our list. SAIC will then finalize the list of technology scenarios and the list of sites to be visited. At each step, SAIC's technical staff identified key words that were used in searching the relevant databases for technologies, companies and organizations that are included in this Information Collection Plan. SAIC's Corporate Technical Resources Acquisition Center (CTRAC) assisted the SAIC technical staff in keyword searches using the databases listed in Table 3.0-2. SAIC's database search capability encompasses several hundred databases in the United States and Europe, including Department of Defense unclassified databases.

Table 3.0-2. Databases Searched by SAIC

Database	Contents
Dialog	Over 400 Commercial Bibliographic Databases
NASA On-Line (RECON)	NASA-Funded Research
Defense Research On-Line Services (DROLS)	Defense-Funded Research
Hyper-Abledata	Products for Sensory and/or Physical Impairments

Dialog is a set of over 400 commercial databases, with the collective scope of an entire library. It includes engineering, medicine, physics and chemistry, education, and business. NASA On-Line covers NASA-sponsored research, and the Defense Technical Information Center's DROLS database covers defense-sponsored research. The Trace Center's Hyper-Abledata database is an excellent compilation of existing products for persons with sensory and/or physical impairments, serving as an indication of what has been done and what companies are on the cutting edge of research for sensory impairments.

Two other sources are vital to information collection. Table 3.0-3 lists some of the journals that SAIC has monitored and/or used databases to search. Table 3.0-4 lists the trade shows that SAIC has attended or will attend for this study. A single trade show can often function as several site visits, and trade shows are an excellent way to experience and evaluate many technologies at one time. Trade shows also offer an opportunity to discuss technologies and scenarios with a wide range of industry representatives, gaining information about existing technologies and products while increasing industry awareness of the needs of the sensory impaired.

Table 3.0-3. Journals SAIC Has Searched

Aerospace Products
Computer Design
Computer Systems News
Defense Electronics
Electronic Business
Electronic Component News
Electronic Design
Electronic Engineering Times
IEEE Proceedings
Instrumentation and Automation News
Military and Aerospace Electronics
PC Week

Table 3.0-4. Trade Shows SAIC Has Attended or Is Scheduled to Attend

National Home Health Care Expo (NHHCE) **November 15-17, 1990** Over 1200 home health care equipment manufacturers and dealers demonstrated their products and services. Table 3.0-5 lists some of the more significant contacts made at the show.
NASA Tech 2000 Show **November 27-28, 1990** Over 150 NASA contractractors and divisions demonstrated and discussed their latest products and research. Table 3.0-6 lists some of the more significant contacts made at the show.
Consumer Electronics Show **January 12-15, 1991** Over 1000 manufacturers presented their latest electronic devices for the home and office. SAIC talked with dozens of manufacturers about the scenarios over the 4-day period.
Comdex Computer Show **June 2-7, 1991** Hundreds of computer and peripheral manufacturers come together in Atlanta to show the state of the art in computing technology.
Armed Forces Communications and Electronics Association (AFCEA) Show **June 4-6, 1991** Over 450 military communications contractors demonstrate and discuss the state of the art in military communications at the Washington D.C. Convention Center.
Federal Microcomputer Conference **August 20-21, 1991** A conference on federal government microcomputers, taking place at the Washington D.C. Convention Center.

BEST COPY AVAILABLE

Table 3.0-5. Contacts Made at the National Home Health Care Expo

Manufacturer/Organization	Technology
American Foundation for Technology Assistance, Inc.	Database of Rehabilitation Products
Bell Regional Companies	Conversion Between Text/Speech/TDD/Modem, Telephone Transmission of Sign Language, Natural Voice Processing
Herrco Enterprises, Inc.	Optical Aids for Low Vision
J.A. Preston Corp.	Voice Synthesis
Kempf	Voice Recognition
Mastervoice	Voice Recognition
McKnight Medical Communications Co.	Directory of Communication Products

Table 3.0-6. Contacts Made at the NASA Tech 2000 Show

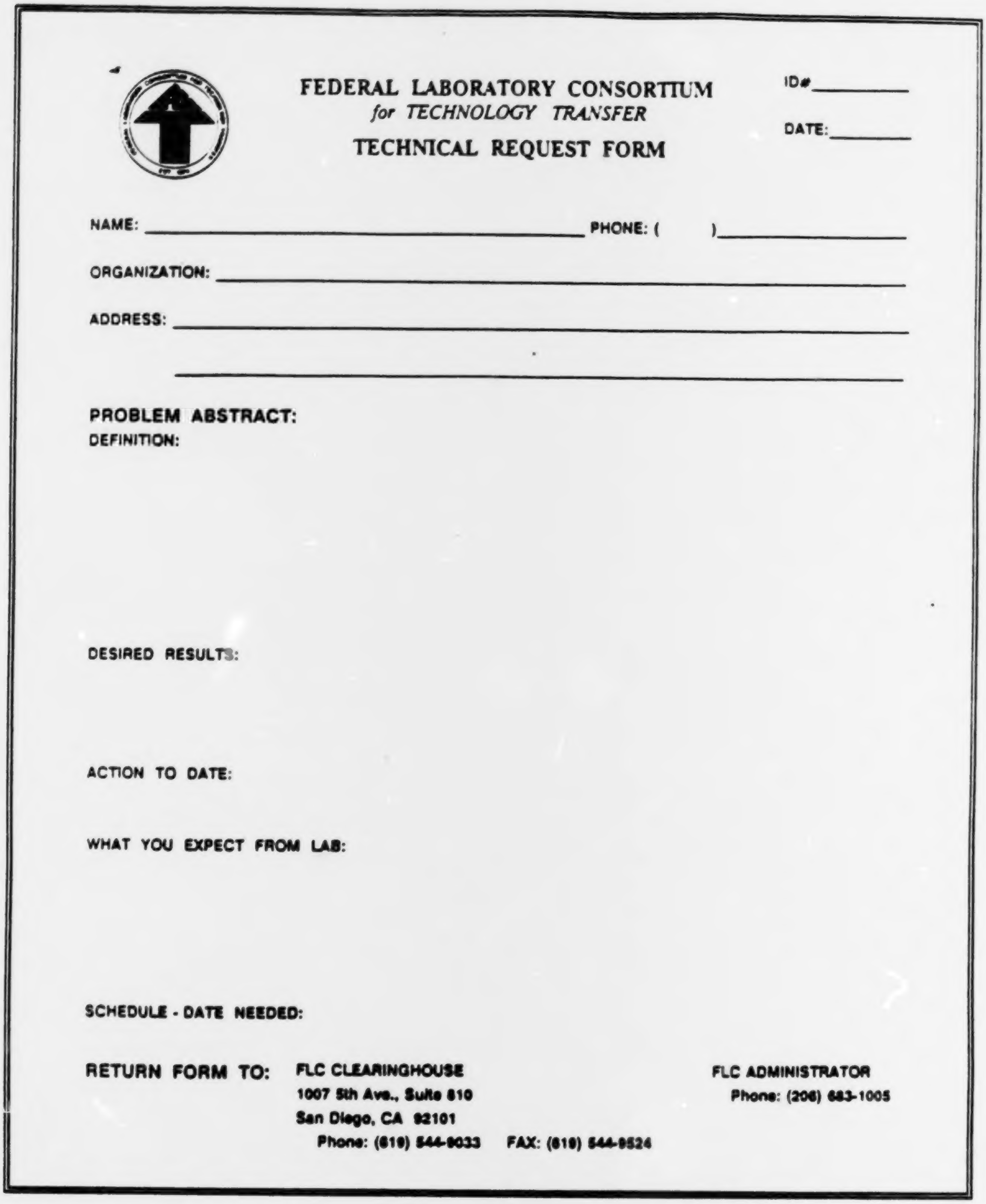
Manufacturer/Organization	Technology
Dolphin Scientific	Voice Recognition
DTI Engineering, Inc.	Direction-Discriminating Hearing Aid
Exos, Inc.	Sign Language Input Device
Federal Laboratory Consortium for Technology Transfer	Night Vision and Other Technologies
Hughes	Infrared Imaging
Infinity Photo-Optical Co.	Vision Enhancement
JR3	Force Sensors
Kodak Federal Systems	Infrared Imaging
NASA	Artificial Reality Research/3-D Audio
Wright Patterson R&D Center	Visual Aids and Audio Applications

Following the database searches, SAIC's technical staff, the Conference Center staff, and CTRAC staff collected articles and papers, and at the same time, identified researchers and organizations to be visited. The information collected has been added to the advanced technology database and will be incorporated into the advanced technology scenarios where applicable. The advanced technology database will be updated throughout the project and will be provided to the Department of Education as part of the final report.

4.0 SITE VISITS

From the data collected, specific researchers and organizations were identified as candidates for site visits. Enclosure 1 contains five lists: Table 1A is the list of industrial sites to be visited or contacted by telephone or mail. Table 1B is the list of NASA sites to be contacted by telephone, mail, or site visit. Table 1C is the list of Federal Laboratory Consortium officers to be contacted by phone and mail, using the form shown in Figure 4.0-1. Table 1D is the list of Department of Defense laboratories to be visited or contacted for information. Finally, Table 1E is a list of the Department of Education Engineering Centers to be contacted. These lists of researchers and organizations are provided to the COTR and Panel of Experts for discussion at the Panel of Experts meeting. Following finalization of the site visit list, a letter will be sent to the selected researchers and organizations to establish a date and time for SAIC and Conference Center personnel to visit. Preliminary contact will then be made by SAIC's technical staff.

In parallel with finalizing the site visit schedule, SAIC will develop a collection schedule and plan, develop a Data Collection Questionnaire, and outline scenarios to be discussed with the sites to be visited. A packet will be prepared for each site to be visited. This packet will include a description of the program, the scenario outline, based on the example of Figure 4.0-2, the questions to be asked, and the expected results of the visit.



FEDERAL LABORATORY CONSORTIUM
for TECHNOLOGY TRANSFER
TECHNICAL REQUEST FORM

ID#________

DATE:________

NAME: ______________________________ PHONE: (   )____________

ORGANIZATION: ____

ADDRESS: ____

____

PROBLEM ABSTRACT:
DEFINITION:

DESIRED RESULTS:

ACTION TO DATE:

WHAT YOU EXPECT FROM LAB:

SCHEDULE - DATE NEEDED:

RETURN FORM TO:

FLC CLEARINGHOUSE
1007 5th Ave., Suite 810
San Diego, CA 92101
Phone: (619) 544-9033 FAX: (619) 544-9524

FLC ADMINISTRATOR
Phone: (206) 683-1005

Figure 4.0-1. Federal Laboratory Consortium Technical Request Form

Target Audiences	Enumerate the target audiences and their potential roles, involvement or contributions. This includes the consumer, technology developer/implementer, producer/manufacturer, Departments of Education, Justice, Defense, Commerce, etc. This is the aggregation of everyone with critical involvement to make this happen.
Technology Application Area	Very brief description of the technology application area, e.g., special displays, captioning, special effects and alternate audio programs on HDTV used for commercial broadcasts, training, and home video for the hearing/visually impaired.
Needs	Brief discussion of the needs, e.g., real-time captioning of voice, special placement of text on the display, special characters colors and sizes, multiple languages, etc., versus the audience. Discuss limitations of the current technology or implementation and how that leaves the need unsatisfied or only partially satisfied.
Technology Application Description	Describe the technology in laymen's terms, i.e., by features, capabilities, limitations and typical/potential applications. Where else would the technology be used, i.e., why is the technology being developed in the first place? What is the maturity of the new technology that will be applied. Who controls or manages the technology, i.e., developers, key patents, centers of excellence, existing/planned investments, commercial/government applications, etc. Describe how the technology would be developed, applied, implemented or modified to provide a capability to the target consumer. What other things must be developed or events occur to ensure the availability for the proposed application. (An obvious example is the development of HDTV along with associated displays, processing, standards, and studio equipment.) Describe the capability that would be provided or extended. Visualize the technology application, e.g., real-time multi-color captioning placed near the speaker, etc. Describe potential synergisms with others' needs, multi-lingual training for the Department of Education, Department of Defense training.
Potential Barriers	Describe potential barriers to the implementation and relate to target audience, e.g., regulatory barriers that require administrative or legislative action, cost barriers or economic incentives, standards that must be established, time or schedule, technology barriers, commercialization barriers, etc.
Est. Costs/ Program Schedule/ Necessary Actions	Lay out skeletal program plan that includes major milestones and overview schedules, required actions and their timeframe, related developments, and estimated costs. Should attempt to show events outside the Department of Education that must take place to ensure successful implementation and completion of the program.

Figure 4.0-2. Sample Scenario Outline

BEST COPY AVAILABLE

An Information Collection Worksheet/Questionnaire, Figure 4.0-3, was developed and will be individualized for each site visit to guide the information collection effort.

1. Site Name:
2. Site Address:
3. Date of Contact:
4. Point of Contact:
5. Sensory Impairment(s):
6. Technology(s):
    a.
    b.
    c.
7. Questions:
    a.
    b.
    c.
8. Applications:
    a.
    b.
    c.
9. Modification(s)/Adaptation(s)
10. Follow-Up Contact:
11. General Comment(s):

Figure 4.0-3. Form for Information Collection Worksheet/Questionnaire

The goals of the information collection process are to:

- Understand specific applications for the advanced technologies.
- Determine the cost to apply the technologies.
- Determine the adaptations of the technologies needed for devices to meet the media access needs of persons with sensory impairments.
- Determine what government supports may be needed to ensure development of the technology into devices to provide information and communication media access.
- Identify the legal aspects of the technology (i.e., patent rights, copyrights and proprietary rights).

To ensure the right questions are asked at the site visit, a list of questions on the technology will be compiled by SAIC's technical staff and reviewed by a subject matter expert from within SAIC or the Panel of Experts. The questions will be provided to the organization/company to be visited. A schedule of visits will be developed based on the recommendations of the Panel of Experts and the COTR's priorities. The list of organizations to be visited will be expanded into a schedule of visits. Visits will be scheduled based on the following considerations:

- Maximize site visit time
- Apply the technologies to specific scenarios
- Maximize technical and marketing staff availability
- Minimize travel costs

5.0 PROCEDURES TO GAIN SITE ACCESS

Key technology centers will be visited and the researchers asked to comment on the outline scenarios that SAIC developed for their technologies related to persons with sensory impairments. The sites to be visited were selected based on the key technologies identified in the database search, recommendations from SAIC's corporate staff, and recommendations from the Panel of Experts.

SAIC's procedures for gaining access to the companies have been to:

- Make telephone contact with the engineering manager responsible for the advanced technology (lowest-level contact possible within the organization).
- Provide written goals and objectives for the visit, if requested.
- Explain how the study can help the organization.
- Provide a two-page description of the Department of Education's program. "Examining Advanced Technologies for Benefits to Persons with Sensory Impairments."
- Identify specific applications for the technology that will be discussed.
- If requested, provide a non-disclosure form stating that SAIC and the government agree not to disclose any proprietary information on a process or device -- except for information available in the open literature -- without written permission of the organization.

6.0 INFORMATION COLLECTION PROCEDURES

Following each visit, a case report will be written and included with the COTR's monthly status report. The visit reports will be used to refine the scenarios and update the technology database. Follow-up visits or telephone interviews will be conducted as necessary to complete the scenario outlines. The scenario outlines will then be used to develop the final scenarios for submission to the Panel of Experts for review in the 14th month of the contract.

7.0 OVERCOMING POTENTIAL BARRIERS

A critical concern in collecting information at each site is the proprietary nature of advanced technology research and development. The protection of devices and applications, to maintain a lead over one's competition, prevents many companies from discussing their applications or technologies with other companies or government agencies prior to the release of a device. SAIC has participated in numerous studies that involved overcoming the problems associated with both the proprietary and classified nature of technology applications.

SAIC is using the following methodology to overcome the proprietary information barrier:

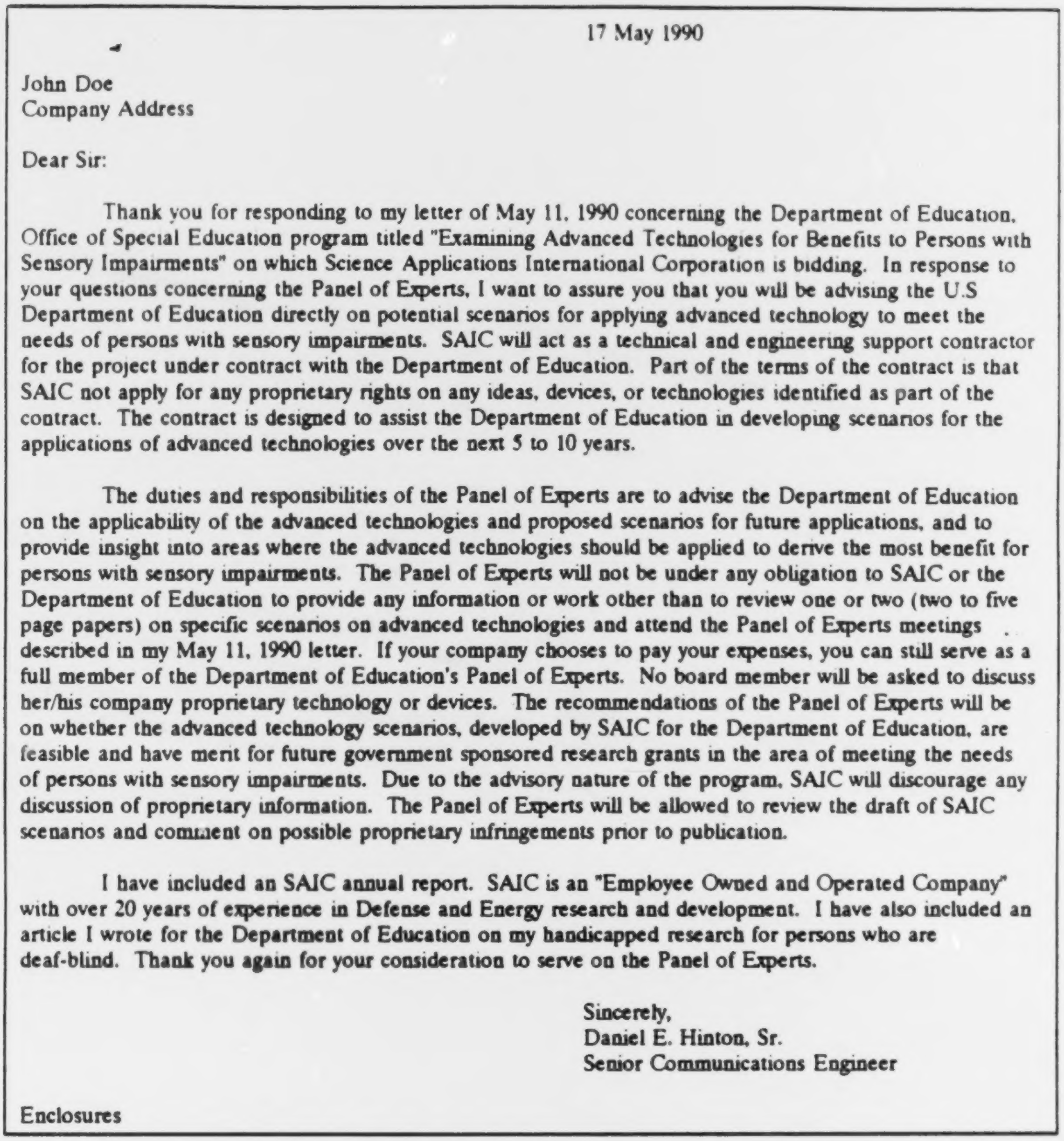
- Identifying potential applications for the technology and discussing these technology applications with the organizations visited.
- Discussing the applications the companies have advertised for their devices.
- Agreeing to sign a non-disclosure agreement, should a company or organization insist on one, to protect technology applications or processes outside the public domain.

SAIC's corporate staff has experience in most advanced technology fields. Therefore, the only discussions necessary are on applications related to understanding how the technologies can be applied to meet the needs of persons with sensory impairments.

An example of a method used to overcome proprietary barriers was how we handled a problem during the formation of the Panel of Experts. One candidate for the Panel of Experts, and his immediate supervisor, were concerned over the possible release of proprietary information. The SAIC Principal Investigator assured that organization's staff that the proprietary work they were doing would not be discussed without prior consent. Figure 7.0-1 is the letter SAIC sent to the company explaining our position and the Department of Education's principal role in the project. The company's executive staff agreed to allow their representative to serve on the Panel of Experts. This personal contact with the technical staff is essential in obtaining the cooperation of both industry and the government departments. The best approach is to contact the technical staff, obtain their support, and then approach management.

For government agencies, such as the Department of Defense, personal contact at the program management level is essential to obtain cooperation in exploring technologies. SAIC's staff has made contact with DoD and other government agencies and identified key government project managers who can make a significant impact on this study. This is the most important step in obtaining high-level DoD cooperation in this study.

To date, SAIC's staff has experienced no barriers to gathering information on advanced technologies. In 10 to 20 cases where we have made contact with engineers or marketing representatives, they have provided technical data sheets, white papers on product applications, and, in one case, a sample of their product. In another case, a small business at the NASA Technology 2000 Show expressed an interest in licencing their technology and providing full technical disclosure. The response to date has been overwhelming. In fact, we are having to be selective in our information gathering to prevent information overload.



17 May 1990

John Doe
Company Address

Dear Sir:

Thank you for responding to my letter of May 11, 1990 concerning the Department of Education, Office of Special Education program titled "Examining Advanced Technologies for Benefits to Persons with Sensory Impairments" on which Science Applications International Corporation is bidding. In response to your questions concerning the Panel of Experts, I want to assure you that you will be advising the U.S Department of Education directly on potential scenarios for applying advanced technology to meet the needs of persons with sensory impairments. SAIC will act as a technical and engineering support contractor for the project under contract with the Department of Education. Part of the terms of the contract is that SAIC not apply for any proprietary rights on any ideas, devices, or technologies identified as part of the contract. The contract is designed to assist the Department of Education in developing scenarios for the applications of advanced technologies over the next 5 to 10 years.

The duties and responsibilities of the Panel of Experts are to advise the Department of Education on the applicability of the advanced technologies and proposed scenarios for future applications, and to provide insight into areas where the advanced technologies should be applied to derive the most benefit for persons with sensory impairments. The Panel of Experts will not be under any obligation to SAIC or the Department of Education to provide any information or work other than to review one or two (two to five page papers) on specific scenarios on advanced technologies and attend the Panel of Experts meetings described in my May 11, 1990 letter. If your company chooses to pay your expenses, you can still serve as a full member of the Department of Education's Panel of Experts. No board member will be asked to discuss her/his company proprietary technology or devices. The recommendations of the Panel of Experts will be on whether the advanced technology scenarios, developed by SAIC for the Department of Education, are feasible and have merit for future government sponsored research grants in the area of meeting the needs of persons with sensory impairments. Due to the advisory nature of the program, SAIC will discourage any discussion of proprietary information. The Panel of Experts will be allowed to review the draft of SAIC scenarios and comment on possible proprietary infringements prior to publication.

I have included an SAIC annual report. SAIC is an "Employee Owned and Operated Company" with over 20 years of experience in Defense and Energy research and development. I have also included an article I wrote for the Department of Education on my handicapped research for persons who are deaf-blind. Thank you again for your consideration to serve on the Panel of Experts.

Sincerely,
Daniel E. Hinton, Sr.
Senior Communications Engineer

Enclosures

Figure 7.0-1. Example Letter to Company on Overcoming Proprietary Barriers

The methodology being used to overcome the classified barrier is to perform an unclassified database literature search to provide a synopsis of various high technologies by keyword. This has proven to be extremely successful. Except for specific applications, the basic research into advanced technologies and devices is generally in the unclassified literature. The criteria being used are based on the potential for applications to be adapted to aid persons with sensory impairments. A team of experts from SAIC, with the proper Department of Defense security clearances, is reviewing the technologies to determine if any of the advanced technology applications or devices could be applied to the needs of persons with sensory impairments. Those advanced technologies which could be applied to the needs of persons with sensory impairments would then be screened to determine if the particular technology or application is classified. If the technology application is not classified then it would be included in the study. Key to this search is the SAIC staff's understanding of the various high technologies being applied by the Department of Defense and how these technologies can be integrated into scenarios showing applications that meet the needs of persons with sensory impairments.

Enclosure 1: Researchers and Organizations to be Contacted

Note: Shading and an asterisk (*) by a location indicate a site visit is planned.

Table 1A. Companies/Researchers to Contact

Scenario Topic	Company/Location	P.O.C./Phone #
Braille (Phase Transition Gels)	Massachusetts Inst. of Tech. * Cambridge, MA	Dr. Toyoichi Tanaka Dr. Atsushi Suzuki (617) 253-4817
Braille/Low Vision/Etc.	TeleSensory Mountain View, CA	Mr. Paolo Siccardo (415) 960-0920
Closed Captioning (PC Video Chip)	EES (Chips and Techs.) Annapolis, MD (chip) New Media Graphics Corp. Billerica, MA (board)	Mr. Craig Davis (301) 269-4234 Mr. Adam Bosnian (508) 663-0666
Computer I/O (AT on a Chip)	ACC Microelectronics Corp. Santa Clara, CA	(408) 727 4356
Computer I/O (Bar Graph/Alarm)	UCE Inc. Norwalk, CT	Dick Borstelmann (203) 838-7509
Computer I/O (Braille Keyboard Interface)	Vetra Systems Corp. * Plainview, NY	Mr. J. DelRossi (516) 454-6469
Computer I/O (Handwriting Input Pen)	Graphics Technology Co. Austin, TX	Joanna Howerton (512) 328-9284
Computer I/O (Large Flat-Panel Displays)	Planar Systems Inc. Beaverton, OR	(503) 690-1100

Computer I/O (LCD Controller Board)	Seiko Instruments USA Torrance, CA	Brian Platt (213) 517-7837
Computer I/O (LCD Controller)	Comm. & Display Sys. Inc. Holtsville, NY	(516) 654-1143
Computer I/O (LCD Controller)	Cybernetic Micro Sys. Inc. San Gregorio, CA	(415) 726-3000
Computer I/O (Miniature Screen Projectors)	Reflection Technology * Waltham, MA	Mr. R. Hoff (617) 890-5905
Computer I/O (Pen-In Windows/Sticky Keys)	Microsoft Corp. Redmond, WA	Mr. Greg Lowney (800) 426-9400
Computer I/O (Pen-Input Computer)	GRiD Systems Corp. Fremont, CA	Mr. Lee Watkins (415) 656-4700 ext. 235
Computer I/O (Pen-Input Computer)	IBM with Go Corp. * Foster City, CA	Ms. Bonnie Albin (415) 345-7400
Computer I/O (Pen-Input Computer)	NCR Corp. with Comm. Intelligence Corp. Menlo Park, CA	Mr. Jake Ward (415) 328-1311
Computer I/O (Pen-Input Computer)	Scenario Inc. Boston, MA	Ms. Judy Bolger (617) 439-6611
Computer I/O (Screen Enlargement)	AI Squared * Atlanta, GA	Mr. D. Weiss (404) 233-7065
Computer I/O (Tiny XT Board)	Ampro Computers Inc. Sunnyvale, CA	Mr. Rick Lehrbaum (408) 522-2100

Computer I/O (UnMouse Mouse Emulator)	MicroTouch Systems, Inc. Wilmington, MA	Mr. Tom Cramer (800) UNM-OUSE
Computer I/O Handheld Touch Screen Comp.	Panasonic Comm. & Sys. Secaucus, NJ	Mr. Marc Schwartz (201) 392-6714
Continuous Speech Recognition	Emerson and Stern * San Diego, CA	Mr. Mark McClusky (216) 331-1261
Electronic Books	Franklin Electronics * South River, NJ	(908) 257-6341
Electronic Books	SelecTronics Inc. Pittsford, NY	(716) 248-3875
Electronic Books	The Reader Project Washington, DC	Mr. Bernie Pobiak Mr. Jon Edelman (202) 667-7323
Emergency Warning (Emergency Vehicle Alarm)	City University of New York New York, NY	Mark Weiss (212) 642-2357
Hearing Aids (Voltage-Controlled Op Amp)	Comlinear Corp. Fort Collins, CO	Karen Cunningham (303) 226-0500
Low Vision (Huge LCD Display)	Seton Name Plate Co. New Haven, CT	Ms. Torie Stillings (800) 451-7084 ext. 528
Low Vision (Video Enhancement)	Wilmer Eye Institute Baltimore, MD	Dr. Robert Massof (301) 955-9653
Media Access	MIT Media Lab * Cambridge, MA	Dr. Wm. Schreiber (617) 253-0300

Modem Access (RS-232/Keyboard Converter)	Vetra Systems Corp. Plainview, NY	(516) 454-6469
Modems (Wireless LAN Card)	NCR Corp. Dayton, OH	Dave Secash (513) 445-4168
Neural Networks	David Sarnoff Research Ctr. * Princeton, NJ	Mr. John Pearson (609) 734-2000
OCR Reader	Praxis (Kurzweil Computer) Washington, DC	Ms. Kathy Conrad (202) 737-0515
Scanners (Hand-Held Type-Independ.)	Caere Corp. Los Gatos, CA	Mr. Dave Hansberry (408) 395-7000
Scanners (Omnidirectional Hand-Held)	NCR Corp. * Ithaca, NY	Mr. Craig Maddox (607) 274-2403
Scanners (Optical Character Reader)	Calera Recognition Systems Santa Clara, CA	Mr. Jim Singleterry (408) 986-8006 ext. 7508
Scanners (Optical Character Reader)	Hewlett Packard Co. Greeley, CO	Mr. Tim Haney (303) 350-4440
Sign Language Recognition	Digital Video Processing * Rockville, MD	Mr. Andrew Girson (301) 670-9282
Telecommunications (FAX Board)	Adtech Micro Systems Inc. Fremont, CA	Cappy Frederick (415) 659-0756
Telecommunications (FAX, etc.)	IDR UniCom Plymouth Meeting, PA	Mr. Mike Yuengling (215) 825-6500
Video Teleconferencing (500:1 Image Compression)	UVC Corp. Irvine, CA	Ken Marsh (714) 261-5336

Video Teleconferencing (Digital Video Interactive)	Intel Corp. Princeton, NJ	Ryan Manepally (609) 936-7636
Video Teleconferencing (Fractal Video Compression)	Iterated Systems Inc. * Norcross, GA	Rick Darby (404) 840-0310
Video Teleconferencing (H.261 Std.)	SGS-Thomson Phoenix, AZ	Thomas Lavallee (602) 867-6279
Video Teleconferencing (Interactive Video/Telephone)	AT&T Microelectronics * Berkeley Heights, NJ	Linda Barducci (201) 771-2000 ext. 2656
Video Teleconferencing (Prog. Compression Chip)	LSI Logic Corp. Milpitas, CA	Peng Ang (408) 954-4880
Video Teleconferencing (Video Compression Chip)	C-Cube Microsystems * San Jose, CA	Katherine Chan (408) 944-6328
Video Teleconferencing (Video Compression Proc.)	Oak Technology Inc. Sunnyvale, CA	Steve Gary (408) 737-0888
Video Teleconferencing Video Signal Processor	Philips Cmpnnts.-Signetics Sunnyvale, CA	Mr. Steve Solari (408) 991-4577
Virtual Environments/ Stereoscopic Displays	Univ. of N.C. Chapel Hill, NC	Henry Fuchs (919) 962-1911 Stephen Pizer (919) 962-1785 Warren Robinett (919) 962-1798
Vision Enhancement (Wide-Angle Hi-Contr. LCD)	Asahi Glass/Optrex Farmington Hills, MI	(313) 471-6220

VLSI Retina	Caltech Pasadena, CA	Prof. Carver Mead (818) 397-2814 Prof. Christof Koch (818) 356-6855
Voice Processing (DSP AD/DA Conv. Chips)	Burr-Brown Corp. Tucson, AZ	John Conlon (800) 548-6132
Voice Processing (FFT Processor Chip)	Plessey Semiconductors Scotts Valley, CA	Steve Brightfield (408) 438-2900
Voice Processing (Telephone Board for PC)	Dialogic Corp. * Parsippany, NJ	(201) 334-8450
Voice Processing & Modems (DSP A/D & D/A Converters)	Atlanta Signal Processors Atlanta, GA	(404) 892-7265
Voice Recog./Pattern Recog. (Neural Networks)	Adaptive Solutions Inc. * Beaverton, OR	Mr. Toby E. Skinner (503) 690-1236 Jennifer Humphrey (619) 691-0890
Voice Recog./Pattern Recog. (Neural Networks)	Bellcore Morristown, NJ	Joshua Alspector (201) 829-4342
Voice Recog./Pattern Recog. (Neural Networks)	Intel & NeuroDynamix Inc. Santa Clara, CA	Mark Holler (Intel) (408) 765-9665
Voice Recognition	Articulate Systems Inc. Cambridge, MA	Ms. Ida McRae (800) 443-7077
Voice Recognition	Dragon Systems, Inc. * Newton, MA	Ms. Flynn (617) 965-5200

Voice Recognition	Kurzweil Appl. Intelligence * Waltham, MA	Mr. John Scarcella (617) 893-5151
Voice Recognition/ Macintosh Icon Recognition	Berkeley Sys. Design, Inc. Berkeley, CA	Mr. Marc Sutton (415) 540-5535
Voice Synthesis (Digital Speech Modules)	Eletech Electronics Anaheim, CA	(714) 385-1707
Voice Synthesis (Pre-Recorded Messages)	Dallas Semiconductor Dallas, TX	Mr. Jim Waldron (214) 450-5322
Voice Synthesis (Voice Output Modules)	Omron Electronics Inc. Schaumberg, IL	Mr. Mark Lewis (708) 843-7900

Table 1B. NASA Technology Utilization Officers to Contact

Scenario Topic	NASA Ctr./Location	P.O.C./Phone #
	Goddard * Greenbelt, MD	Mr. Donald Friedman (301) 286-6242
	Langley * Hampton, VA	Mr. Joe Mathis (804) 864-2490
Data Compression/ Neural Networks/ Speech Encoders/ VLSI Retinas	JPL * Pasadena, CA	Ed Beckenbach (818) 354-3464
Emergency Warning (Emergency Vehicle Alarm)	Langley * Hampton, VA	Dr. Bucky Holmes (804) 864-4649
Emergency Warning (Emergency Vehicle Alarm)	Marshall MSFC, AL	James Currie (205) 544-3524
I/O Devices/ Virtual Environment	Ames Moffett Field, CA	Dr. Michael McGreevy (415) 604-5784 Scott Fisher
Image Processing	Stennis SSC, MS	Dr. Doug Rickman (601) 688-1920
Low Vision	Ames Moffett Field, CA	Dr. Jim Larimer (415) 604-5185
Neural Nets/ Sign Language Translation/ Video Manipulation/ Voice Recognition	Johnson * Houston, TX	Dean Glenn (713) 283-5325

Neural Networks	JPL Pasadena, CA	Anil Thakoor (818) 354-5557
Neural Networks/ Voice Recognition	Johnson * Houston, TX	Robert Savely James Villareal (713) 483-8105
Sign Language	Johnson Houston, TX	Robert Savely James Villareal (713) 483-8105
Video Compression	Lewis Cleveland, OH	Wayne Whyte (216) 433-3482 Mary Jo Shalkhauser (216) 433-3455
Video/Image Warping (ARM)	Johnson Houston, TX	Dr. Richard Juday (713) 483-1486
Vision Enhancement (For Maculopathies)	JPL * Pasadena, CA	Dr. Teri Lawton (818) 354-4257

Table 1C. Federal Laboratory Consortium Contacts

Scenario Topic	Federal Lab Consortium/Location	P.O.C./Phone #
	Mid-Atlantic Regional Coordinator Washington, DC	Mr. Nick Montanarelli (202) 653-1442
	Washington, DC Rep. Washington, DC	Dr. Beverly Berger (202) 331-4220

Table 1D. Defense Department Laboratory Contacts

Scenario Topic	Defense Department Lab/Location	P.O.C./Phone #
Image Intensifiers Infrared Imaging	CECOM Night Vis./Electro-Opt. Ctr. * Fort Belvoir, VA	Mr. B. Freeman, Sr. Sci. (703) 665-5508
Multiple Scenarios	DARPA Defense Mfg. Office * Arlington, VA	Dr. Michael J. Kelly (703) 697-6507
Voice Processing Noise Reduction	Rome Air Dev. Ctr. Audio Lab * Griffiss Air Force Base, NY	Mr. Henry Bush (315) 330-7052

Table 1E. Department of Education Engineering Center Contacts

Scenario Topic or R.E.C. Specialty	R.E.C. Name/Location	P.O.C./Phone #
Access to Computers & Electronic Equip.	Trace Center, U. of Wisconsin Madison, WI	Dr. Gregg Vanderheiden (608) 262-3822
Augmentative Comm. (Visual Telephone)	University of Delaware Newark, DE	Dr. Richard Foulds (302) 651-6830
Computer I/O (Blind Keybrd./Synth.)	WesTest Engineering Corp. Bountiful, UT	Mr. James S. Lynds (801) 298-7100
Computer I/O (Vis. Imp. Scrn. Acc.)	Automated Functions, Inc. Olney, MD	Mr. Ronald A. Morford (301) 774-0114
Computer I/O (Voice Synthesis)	Dynamic Industries Corp. Deer Park, NY	Mr. Les Schonbrun (516) 667-0448
Emergency Warning (Sound Recognizer)	Applied Concepts Corporation Winchester, VA	Mr. Richard I. Johnson (703) 722-1070
Evaluation of Rehabilitation Tech.	Natl. Rehabilitation Hospital Washington, DC	Ms. Jan Galvin (202) 877-1932
Hearing Aids (Visual/Tactile)	Univ. of Miami, Mailman Ctr. Miami, FL	Dr. Rebecca Eilers (305) 547-6350
Real-Time Captioning (Court Steno to Text)	CADSA, Inc. Webster, TX	Dr. Bartus Batson (713) 338-2691
Real-Time Captioning (Court Steno to Text)	Netrologic San Diego, CA	Mr. James R. Johnson (619) 587-0970

Real-Time Captioning (Court Steno to Text)	Virgus Computer Systems Seattle, WA	Dr. Paul A. Knaplund (413) 736-7299
Real-Time Captioning (Steno-Speech Feas.)	Adv. Technologies Concepts Reston, VA	Mr. Franklin D. Smith (703) 450-7847
Rehabilitation Technology Transfer	Electronic Ind. Foundation * Washington, DC	Dr. Lawrence Scadden (202) 955-5823
Study of Adapt. Uses of Tech. by Disabled	Mr. James C. Dickson Washington, DC	Mr. James C. Dickson (202) 832-6564
TDD Modem Access (w/Text-to-Speech)	Integrated Microcomputer Sys. Rockville, MD	Dr. C. Eric Kirkland (301) 948-4790
Tech. Aids for the Deaf/Hearing Impaired	The Lexington Center, Inc. * Jackson Heights, NY	Dr. Harry Levitt (718) 899-8800
Visual Impairments (Low Vision) Hearing Aids (Interaural Delay)	Smith-Kettlewell Foundation San Francisco, CA	Dr. Arthur Jampolsky (415) 561-1630 Dr. Helen J. Simon (415) 561-1681

APPENDIX C

TEN YEAR DEVELOPMENT PLAN

1.0 GENERAL

This development plan covers the program "Examining Advanced Technologies for Benefits to Persons with Sensory Impairments" in which SAIC developed ten scenarios for the Department of Education Office of Special Education Programs. Attached is a ten year development program plan time line developed under this study. There were three categories of technologic improvements considered for the visually impaired, hearing impaired, and the visual and/or hearing impaired. In the latter category no further research and development by the Department of Education was recommended for portable power supplies because the commercial marketplace and consumers will ensure development in this area.

As a first step to utilizing this study, the Department of Education's program manager for advanced technology research and development should meet with the architects of this study to help him formulate a detailed five to six year plan using the recommendations of this study as a basis. This will help the Department of Education in funding its goals on a priority basis. It will also prevent duplication of effort in similar areas. For example, speech synthesis applies to both the Character Readers for Display scenario and the Input/Output Devices scenario, but the same speech synthesis technology would apply to both applications.

It is also recommended that another study be performed in another three to five years to examine advanced technologies for benefits to persons with sensory impairments. As a result of this future study, another five to six year plan should be formulated, based on technology that appears in the interim. Thus, with the state of technology changing as rapidly as it is the planning cycle will not be overcome by events which make it outdated.

2.0 TEN YEAR PLAN

The attached ten year plan is an upper level time line of the three technology areas covered in this study. The first five years are the most detailed since there is more certainty about the near future. The following five years is less detailed since the future will undoubtedly contain new and miraculous discoveries and inventions which will profoundly affect how the sensory impaired population can benefit from technology. Also, Braille devices and descriptive video tend to require research beyond the five year time frame, as do Visible Light Spectrum Manipulation and Character Readers for Displays. Speech Recognition Systems and Video Teleconferencing are already becoming mainstream consumer products and services, and the other technological scenarios lend themselves to more immediate solutions and developments and/or replacement with new technologies as they develop.

Braille devices technology, unlike some of the other categories, will require major spending initiatives. This is a result of the level of investment that the technology developers will require. It is estimated that approximately $1 million for the next five years will be required to invest in these Braille device technologies. It is hoped that a cooperative effort between NASA and the Department. of Defense would be able to fund the small actuator development as a first step to developing advanced Braille devices.

Voice recognition technology is another area that will require large investments by the Government to encourage development of input/output devices for persons with hearing and/or visual impairment. It is estimated that a funding level on the order of magnitude of $1 million per year for the next five years would be required. This funding could be awarded to universities to fund research on interpreter service devices, telephone (TDD) relay, and closed captioning speech recognition systems.

Descriptive Video is also a technology area which will require more than small grants and SBIR funding to entice research and development that will benefit the visually impaired. It is estimated that approximately $200,000 will be required the first year and $300,000 to $500,000 the next two to five years to encourage DV development.

3.0 NEAR TERM PLANNING (FIVE YEAR DETAILED PLAN)

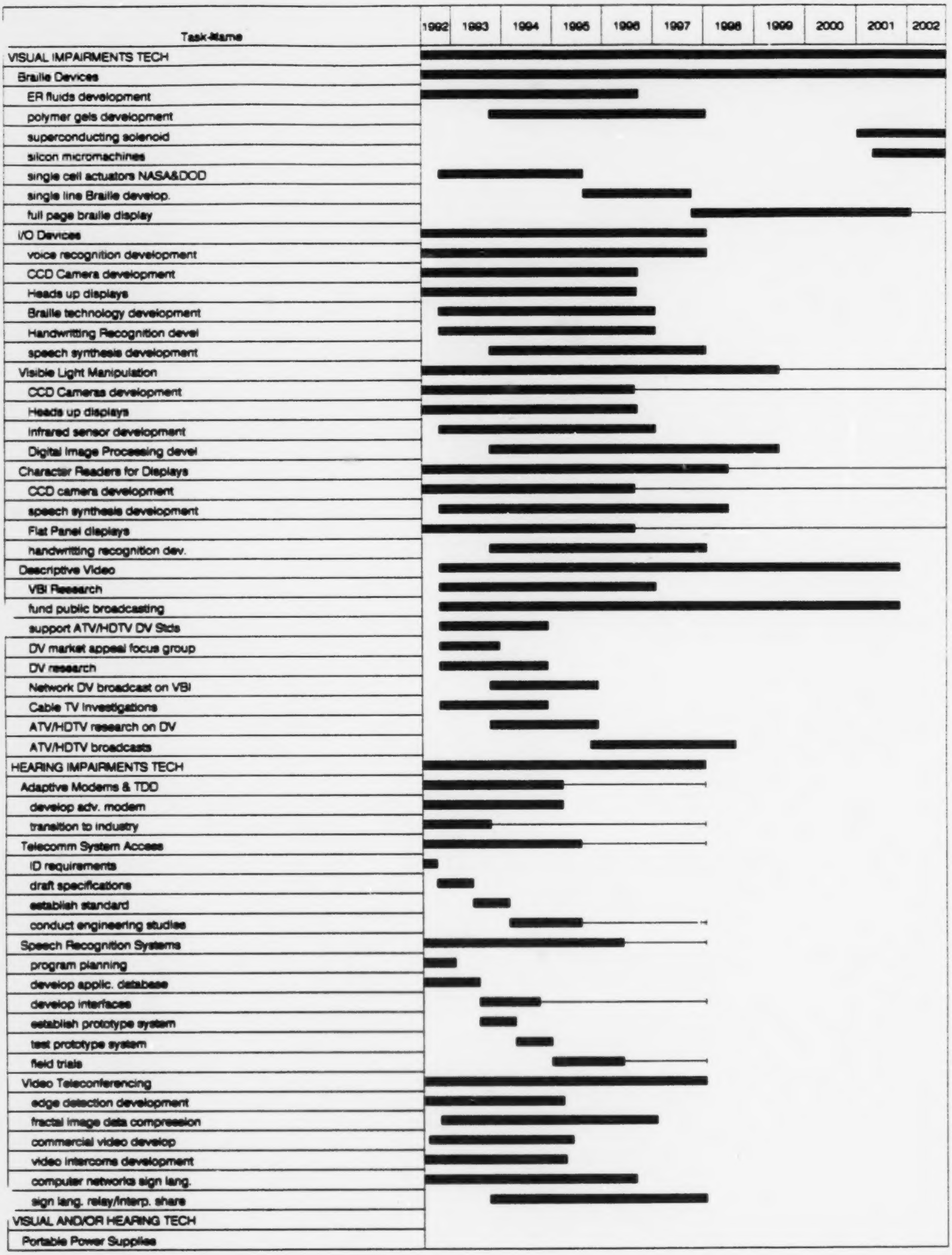
Attached for reference is a time line plan for developing a detailed five year plan. Prior to beginning the task of developing a detailed five year research and development plan for the OSEP. It should be understood that plan development must be approached with help from others (i.e.,panel of experts, task force, contractors, etc.).

The first step is to review documentation which may be pertinent to the next five year cycle, such as the SAIC study scenarios, or at least their conclusions and recommendations. This will provide a background as to what may be in the offing. Step two would be establishing the goals of the research and development program that the Department of Education Office of Special Education Programs (OSEP) might want to foster and accomplish in the near future. This data could be gathered from OSEP personnel as well as the panels own insight. Step three would be for the special group to develop a draft five year plan, which includes the different technological areas as well as the time frame for their studies, grants, and assistance to be administered. Administering the resulting plan would then require staffing by the Government or appointed body (maybe a panel of experts). Step four would be the development of a budget in which to administer the five year plan. This part of the exercise would add some realism to the detailed plan provide budget cycle planning and programming information. The fifth and final step would be to put the final touches on the detailed plan and obtain approval for its inception.

4.0 FUTURE PLANNING

As discussed earlier, another 18 to 24 month study effort should take place in the next three to five years (1995-1997) to determine what technology may be applied to benefit persons with sensory impairments so that a subsequent detailed five year plan may be instituted for that time period (1997-2001). This would again be the basis for convening a task force to develop a detailed five year plan so that research and development funding can be applied in a coherent and responsible manner. Especially since these development efforts are funded with tax dollars, it is essential to coordinate those resources to get the most research and development for the limited resources available. That way funds can be parceled out based on careful study and analysis that includes the broader technological issues, putting trends and emotional issues into perspective with cost and feasibility.

TEN YEAR DEVELOPMENT PLAN



FIVE YEAR DEVELOPMENT PLAN

Task Name	Duration	Start	End	1992			
				Jun	Jul	Aug	Sep
Detailed 5 Year Planning	3.5 m	6/1/92	9/16/92	▲——	———	———	——▲
Review Documents	0.50 m	6/1/92	6/15/92	▲—▲			
Establish Goals	0.50 m	6/15/92	6/30/92	▲—▲			
Develop Draft Plan	1.00 m	6/30/92	7/30/92	▲——	——▲		
Review and Discuss Plan	0.50 m	7/30/92	8/14/92		▲——	——▲	
Develop Budget	0.50 m	8/14/92	8/31/92			▲—▲	
Finalize Plan	0.50 m	8/31/92	9/16/92			▲——	——▲

APPENDIX D

SCENARIOS

BRAILLE DEVICES AND TECHNIQUES TO ALLOW MEDIA ACCESS

MARCH 1992

Prepared by

Daniel E. Hinton, Sr., Principal Investigator
and
Charles Connolly

SCIENCE APPLICATIONS INTERNATIONAL CORPORATION
3701 N. Fairfax Drive, Suite 1001
Arlington, VA 22203
(703) 351-7755

1.0 SCENARIO

Braille Devices and Techniques to Allow Media Access.

2.0 CATEGORY OF IMPAIRMENTS

Persons with vision impairments.

3.0 TARGET AUDIENCE

Consumers with Vision Impairments. Persons with vision impairments will benefit from enhanced access to media information services and computer systems. This scenario on advanced materials and technology for implementing Braille provides a means to disseminate information to consumers with vision impairments. In particular, it provides a better understanding of the technology available to produce Braille over the next three to five years.

Policy makers, including national representatives, Government department heads, and special interest organizations. Policy makers can use this scenario to better understand the issues related to media access for persons with vision impairments. In addition it provides a point of departure for policy makers to understand how advanced technology legislative or regulatory funding priorities within Government programs can accelerate Braille output device development.

Researchers and Developers. This group will benefit through a better understanding of the needs of persons with vision impairments and specifically their printed media communications needs. Understanding media access requirements will assist researchers and developers in designing Braille media access functions into their future products to meet the needs of persons with vision impairments.

Manufacturers. Manufacturers will benefit through a better understanding of Braille device requirements, the potential market size and the existing Federal Government requirements for media access for persons with vision impairments which can be met by adding a Braille capability to their systems.

4.0 THE TECHNOLOGY

Louis Braille published a dot system of Braille in 1829 based on a "cell" of six dots. He defined the alphabet, punctuation marks, numerals, and later a system for music using the 63 possible dot arrangements. Braille is read by running a finger over a character and sensing the raised dot pattern. Braille output devices in use today include a stylus on a pocket-sized metal or plastic slate (analogous to a pencil and clipboard), Braille writers like the Perkins Brailler (analogous to typewriters), and computer Braille devices. Printed Braille can be stamped on both sides of a page in a process called interpointing. This process saves paper and reduces the size of Braille books.

The nominal specifications for Braille dot, Braille cell and Braille page dimensions are set by the National Library Service for the Blind and Physically Handicapped (NLS). The NLS certifies all Braille transcribers sponsored by the Library of Congress based on these specifications:

Braille dots:
- Height for paper Braille, 0.019 inches, uniform within transcription;
- Base diameter, 0.057 inches;

Braille cell:
- Center-to-center distance between dots, 0.092 inches;

Corresponding dots of adjacent Braille cells:
- Horizontal separation, 0.245 inches;
- Vertical (down page) separation, 0.400 inches;

Braille page:
- Standard size, 11.5 inches wide by 11 inches high.
- Minimum margin for binding side, 1 inch;
- Minimum margin for other sides, 0.5 inch;
- Minimum weight of paper, 80-pound;
- Paper must be thick enough so that, at worst, 10% of dots break the paper surface, but thin enough to permit uniform dots of proper height.

The Perkins Brailler, made by the Howe Press of the Perkins School for the Blind, Watertown, Massachusetts, is a machine for embossing characters on paper. It is widely regarded as the standard for quality within the industry and has been used for over 100 years. It is capable of embossing 25 lines of 40 characters each, which is the page layout implied by the NLS standards. The term "Perkins" is used almost generically to refer to all Braille machines.

It should be noted that even though 11x11.5 inch paper is standard, many rely on 8.5x11 inch paper because it works well with a slate and stylus.

For paperless Braille, approximately 20 grams of force at 0.010 inches displacement, and 0.020-0.030 inches displacement without opposing force, may be a useful guideline for acceptable feel. Different technologies have different force-displacement characteristics which Braille-literate people must evaluate on a case-by-case basis.

The Braille Authority of North America is the committee that sets standards for Braille code in the U.S., and all sanctioned Braille code is based on a 6-dot Braille cell: 2 columns of 3 dots each. Figure 1 shows the Braille alphabet. Nemeth Code, which is the standard Braille notation for mathematics, Computer Braille Code, and Textbook Braille

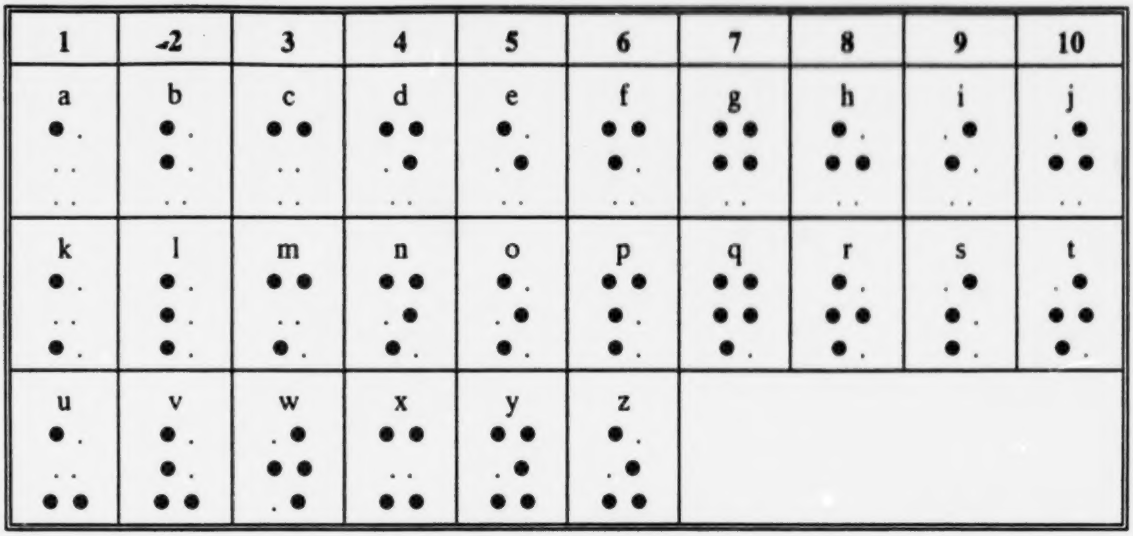


Figure 1. Grade 1 Braille Alphabet

are all based on 6-dot Braille cells as is the literary Braille used for mainstream text translation.

Some paperless Braille cells are produced in the U.S. with 8 dots per cell--2 columns of 4 dots each--but these are for compatibility in the European market. In Europe, the extra two dots are used to represent upper case letters and computer characters: control characters and extended ASCII characters. Eight-dot Braille cells could be adopted as the standard for computers in the U.S., but that seems unlikely for five reasons:

1. Six-dot Braille has a long and successful history in the U.S., and the cost of replacing Braille printers and paperless Braille displays in a short time would be prohibitive.

2. More lines of 6-dot Braille fit on a page than 8-dot Braille, and Braille already takes up several pages per printed page. An embossed Braille document takes about 15 to 20 times the volume of the same document printed, making it even less likely that 8-dot cells would be adopted for Braille books. This is based on the fact that it takes approximately three Braille pages per standard print page and a Braille page is 11x11.5 inches vs 8.5x11 inches for a standard text page. Also, a Braille page is 3 to 6 times thicker than a standard text page.

3. No single standard has emerged for special computer characters in 8-dot Braille.

4. One third more dots per cell would make an 8-dot Braille output device considerably more expensive than a 6-dot output device; the cost per dot generally dominates the total cost of Braille displays.

5. The "War of the Dots," which ended in 1918 with the choice of modified French Braille notation over American Braille and New York Point, has made Braille experts extremely cautious about making major changes in Braille notation.

5.0 STATEMENT OF THE PROBLEM

Persons with visual impairments have limited real-time access to computer information because existing Braille output devices are expensive and can only display 20-80 characters at a time. In the U.S., voice synthesis devices are used by more visually impaired Americans than paperless Braille devices due to their lower cost. Paperless Braille displays are more common in Europe, where the Government generally pays for displays. Affordable paperless Braille is needed because voice synthesis does not allow the user to quickly review material as it appears on the monitor or printed page, including its format and structure. With the advent of large CD-ROMs with database libraries containing millions of print characters, and the increasing availability of information accessible by computer, persons with visual impairments need Braille displays that allow them equal access to the text displayed for sighted persons on the computer monitor. The best Braille displays now available limit persons with severe vision impairments to a single line of 20, 40 or 80 Braille characters. This makes it difficult to scan through text files and look for headings or jump from paragraph to paragraph. There is an urgent need for larger Braille displays to allow persons with vision impairments text access capability equivalent to that of sighted persons.

Several factors influence the demand for Braille displays: the rate of Braille literacy is low among persons with vision impairments in the U.S., perhaps 20 percent. This is because, in part, visual impairments often set in with advancing age when it is more difficult to learn Braille. Most legally blind Americans are elderly. Also, age can adversely affect hearing, so there are older Braille-literate Americans who cannot use voice synthesis technology. The segment of the population with deaf-blindness with little or no residual hearing, regardless of age, also cannot benefit from voice synthesis technology. Some people cannot use Braille because they have reduced tactile sensitivity, as with diabetes, age, and occupations that callous hands. Overall, the largest demand for paperless Braille in conjunction with computers comes from people who can use voice synthesis technology but, because of the need to study, review and edit text, need to use paperless Braille.

Many people with vision impairments want the capability to produce computer-driven Braille displays containing 3 or 4 horizontal lines of 80 Braille characters each. Others want 3 or 4 lines of 40-42 Braille characters. Many persons with vision impairments would be satisfied with a refreshable Braille display that simulates the 25-line Perkins

Brailler page. However, size, weight, power, reliability and cost per unit will determine the maximum Braille page size.

Researchers should focus their attention on identifying fresh approaches to producing the dots required to form the Braille characters within the space limitations imposed by the Braille specifications listed in Section 4.0.

According to Noel Runyon, an engineer at Personal Data Systems and a Braille user, the critical factors that are easiest to overlook in the design of a full-page Braille display include:

1. Speed. Most reading is skimming, not sequential, cover-to-cover reading. Also, people can learn to read Braille as fast as sighted people read print.
2. Navigation. If display updates cannot occur in the blink of an eye, it is important to be able to "point" to a part of the display, evaluate it, and go to another page without waiting for the entire display to update, because everyone needs to flip through pages. Single characters must be individually addressable, and readers need a feel for where they are on a page.
3. Cursor location is critical on a computer display.
4. Application-specific devices are too restrictive to meet the broader communication needs of Braille users. For example, sequential output devices are awkward for most types of reading, whether their output is Braille or speech.
5. Humble things like dust can render laboratory successes almost useless in real homes and offices.
6. Graphics capability is a major justification for the use of a full-page display rather than a smaller display.
7. Battery power is highly desirable.
8. Noise is an important factor, especially in offices, libraries, and other public places.
9. Cost can make the difference between a device that is evolutionary and a device that is revolutionary.
10. Elderly and pre-employment-age people have generally received the least attention when developing new Braille technology, so they tended to be left out of the decisions that led to existing devices.

6.0 THE DEPARTMENT OF EDUCATION'S PRESENT COMMITMENT AND INVESTMENT

The primary reason that Braille media access is a priority is because approximately 100,000 Americans with vision impairments use Braille for written communication. According to the 1988 National Health Interview Survey, 600,000 Americans between the ages of 18 and 69 have blindness or visual impairments severe enough to limit their employment opportunities, and that number rises sharply with age. This is an indication of the size of the population who could potentially benefit from Braille literacy. Although the number of visually impaired people under 18 is relatively small, they can learn Braille most easily and use it for the rest of their lives, thus they can gain the most from Braille literacy.

The Department of Education, and its predecessor, the Department of Health Education and Welfare (HEW), have funded Braille device research and development over the past 20 years. With the advent of personal computers in 1975, HEW began to fund research and development of computer Braille output devices such as the TeleBrailler, and MicroBrailler. Currently, the development of Braille capability is a stated research priority of the Department of Education as follows:

- The Electronic Industries Foundation (EIF) Rehabilitation Engineering Center's Technology Needs Assessment Paper, "An Inexpensive Refreshable Braille Display," points out a need for a "low-cost, reliable paperless Braille display mechanism." That report follows up on the recommendations of the National Workshop on Rehabilitation Technology, sponsored by EIF and the National Institute on Disabilities and Rehabilitation Research (NIDRR). The Workshop recommended making "information processing technology for access to print graphics, including computer access" the top technology priority for visual impairments.

- Several of the funding criteria of the Department of Education's National Institute on Disability and Rehabilitation Research (NIDRR) are directed at the high unemployment and underemployment rate of persons with vision impairments and severely visually impaired populations. Most severely visually impaired Americans are unemployed. Larger and more affordable Braille displays would improve the educational outlook of blind individuals, promote Braille literacy, and improve employment opportunities and job retention among the Braille literate. Another stated priority, advanced training for the blind and visually impaired at the pre- and post-doctoral levels, and in research, would benefit greatly from improved Braille display technology.

- The Panel of Experts for the Department of Education program sponsoring this study consists of experts from industry and Government, including members of the sensory-impaired community. Their consensus opinion was that developing a larger Braille display is the highest priority for persons with visual impairments.

- One of the Department of Education's 1991 Small Business Innovative Research (SBIR) Program Research Topics is to develop or adapt communication devices for young children who are blind or deaf-blind. An affordable Braille display could be used for games that would help young children develop the skills needed to read and write Braille. A Braille display would also be of some use for tactile graphics, though an evenly spaced array of dots based on the same technology might be better.
- The Department of Education's NIDRR Program Directory, FY89, lists the Smith-Kettlewell Rehabilitation Engineering Center, among many other tasks, as testing, developing, and/or evaluating a Braille display technology.

7.0 ACCESS TO COMMUNICATIONS MEDIA

Many federal, state, and local laws which influence access for persons with visual impairments. The most important single law related to access for persons who are vision impaired is Public Law 101-336, enacted July 26, 1990. Better known as the Americans with Disabilities Act (ADA), this law has broad implications for all disabled Americans and establishes the objective of providing access to persons with disabilities to physical and electronic facilities and media.

The other law that impacts technology for persons with visual impairments is Public Law 100-407-AUG. 19, 1988 titled "Technology-Related Assistance for Individuals with Disabilities Act of 1988." Also known as the Tech Act, this law established a comprehensive program to provide for technology access to persons with disabilities. The law defines assistive technology devices:

> "Assistive technology devices means any item, piece of equipment, or product system, whether acquired commercially off the shelf, modified, or customized, that is used to increase, maintain, or improve functional capabilities of individuals with disabilities."

Braille technology clearly meets this definition for persons with vision impairments and should be exploited to increase the ability of persons with vision impairments to obtain access to printed media. Within the findings and purpose of this law, Braille technology can provide persons with vision impairments with opportunities to:

- exert greater control over their own lives by making literacy possible;
- participate in and contribute more fully to activities in their home, school, and work environments, and in their communities;

- interact with nondisabled individuals; and
- otherwise benefit from opportunities that are taken for granted by individuals who do not have disabilities.

8.0 POTENTIAL ACCESS IMPROVEMENTS WITH ADVANCED BRAILLE TECHNOLOGY

Table 1 shows a sampling of the Braille technology currently available. The base price of adding paperless Braille to a computer is now about $5000. This high cost forces many persons with visual impairments in the U.S. to use voice synthesizers which costs about $1000. Braille embossers starting at approximately $1700 for the Braille Blazer, cost about three times that cost of text printers used by the sighted population.

Advanced Braille technology offers persons with visual impairments the potential for dramatic improvements in access to books and periodicals stored in computer-readable form or scanned. For example, at least 800 titles are already available on CD-ROM and that number will probably increase rapidly in the years to come. Another important access improvement would be to computer-based telecommunications, including databases, electronic mail systems, computer bulletin board systems and mail order systems, all of which generally consider a computer screen as a single unit. One-line paperless Braille displays have been a cost compromise when compared to the speed and agility that a full-screen display could offer.

It is often desirable to skim text for relevant information, whether that text is a computer's display, magazine or newspaper article, or book. When skimming, the field of the display needs to be as large as possible. The only practical alternative is Braille paper output, but relying on a Braille paper printer (priced for individual use) is slow and paper-intensive. A multiple-line paperless Braille display offers tremendous improvements in skimming speed and effectiveness over existing Braille printers and single-line displays. It also would have a great impact on the ability of persons with vision impairments to do research and academic study, which often requires reading and rereading information.

9.0 ADVANCED BRAILLE TECHNOLOGIES

There are two major approaches to producing paperless Braille. The simplest approach is to apply constant power to keep each dot raised or lowered, but many of the technologies used to move dots require a substantial amount of power (50 to 100 milliwatts per cell). An analysis of the power available to a full page Braille display provides insight into the power that can be allocated to each Braille dot/cell. In older houses, standard electrical outlet can provide about 1200 watts of power. About 250 watts of that must be allocated to the computer controlling the display, leaving 950 watts for the Braille display. Assuming the display's power supply is 50% efficient, that leaves only 475 watts of power in the form the display can use. An 80-cell display with 6 dots per cell can allocate almost 1 watt per dot; 8 dots per cell lowers that to about 0.75 watts per dot. A standard Braille

Table 1. A Sampling of Existing Braille Products
Note: Prices range from 1989-1991 so they may not be comparable.

Brand Name	Manufacturer	Price	System	Description
HARD COPY				
Perkins Brailler	Perkins School for the Blind	$395-$730	None	Braille Writers, Manual and Electric
Mountbatten	HumanWare Inc.	$2595-$3170	None	Braille Writer, Electronic
Index Braille Embossers	HumanWare Inc.	$2895-$16,900	IBM	Braille Embosser
Braillo 90	Braillo Norway AS	$5795	IBM	Braille Embosser
Braillo 200	Braillo Norway AS	$39,995	IBM	Braille Embosser
Braillo 400 S	Braillo Norway AS	$78,995	IBM	Braille Embosser
Romeo Brailler	Enabling Technologies Company	$2695-$3450	All	Braille Embosser
Marathon Brailler	Enabling Technologies Company	$11,500	All	Braille Embosser
TED-600 Text Embossing Device	Enabling Technologies Company	$37,500	All	Braille Embosser
Braille Blazer	Blazie Engineering	$1695	All	Braille Embosser
ATC/Resus 214 Printer	American Thermoform Corporation	$15,995	All	Braille Embosser
Versapoint-40 Braille Embosser	Telesensory Corporation	$3795	All	Braille Embosser/Translator
Ohtsuki BT-5000 Braille/Print Printer	American Thermoform Corporation	$5180	IBM Apple	Braille Embosser/Printer
Duran Dots-40	Arts Computer Products Inc.	$710-$1510	IBM	Adapter to Convert Brother HR-40 Daisy Wheel Printer for Braille Printing
Stereo Copy Developing Machine	Matsumoto Kosan Company	$6250	None	Braille Copier
Thermoform Duplicators for Braille	American Thermoform Corporation	$1750-$2895	None	Braille Copiers
Plate Embossing Device PED-30	Enabling Technologies Company	$62,500	None	Braille Plate Embosser for Printing Houses
TACTILE READING SYSTEM				
Optacon II	Telesensory Corp.	$3495-$3995	All	Portable Tactile Reading System
InTouch	Telesensory Corp.	$395	Mac	Optacon II Accessory Software for Mouse Access
Optacon PC	Telesensory Corp.	$395	IBM	Optacon II Accessory Software/Hardware for Mouse Access
ONE-LINE BRAILLE DISPLAYS				
Braille Display Processor	Telesensory Corp.	$3695	IBM Apple	Paperless Braille 20 Cells

BEST COPY AVAILABLE

Braille Display Processor BDP 21	Telesensory Corp.	$3695	IBM	Paperless Braille/Translator 20 Cells
Braille Display Processor BDP 20	Telesensory Corp.	$3695	Apple	Paperless Braille/Translator 20 Cells
Braille Interface Terminal	Telesensory Corp.	$3995	IBM	Paperless Braille 20 Cells
Navigator	Telesensory Corp.	$3,995-$14,995	IBM	Paperless Braille 20,40,80 Cells
VersaBraille II+	Telesensory Corp.	$5995	IBM	Portable Paperless Braille 20 Cells
KeyBraille	HumanWare Inc.	$5025-$7025	Toshiba	Paperless Braille 20,40 Cells
Alva	HumanWare Inc.	$8,995-$14,495	IBM	Paperless Braille 40,80 Cells
Braillex IB80	Index Inc.	$14,495	IBM	Paperless Braille: 80 Cells
New Ability Brailler	Densttron Corp.	$2995	IBM	Paperless Braille (Soft Braille) 40 Cells
BRAILLE NOTES/COMPUTERS				
Notex	Index Inc.	$5800-$7900	IBM	Portable Braille Notetaking Device/Computer with 20- or 40-Cell Paperless Braille
Personal Touch	Blazie Engineering	$5500	All	Portable Braille Notetaking Device/Computer with 20-Cell Paperless Braille
Braille 'n Speak	Blazie Engineering	$905	All	Portable Braille Notetaking Device/Translator
SpeakSys	Blazie Engineering	$149	IBM	Braille 'n Speak Interface
PocketBraille	American Printing House for the Blind	$905	All	Portable Braille Notetaking Device/Word Processor
Eureka A4	Robotron Access Products Inc.	$2595	IBM	Portable Talking Computer with Braille Keyboard
Nomad	Syntha Voice Computers Inc.	$2295	All	Portable Talking Computer with Braille Keyboard Option
FOR THE DEAF-BLIND POPULATION				
AFB Tellatouch, MS 170	American Foundation for the Blind	$595	None	Typewriter Keyboard Controlling a Paperless Braille Cell for 1-Way 1-on-1 Communication
DiaLogos	Finnish Central Association of the Visually Handicapped		None	Braille Keyboard with Six Paperless Braille Cells Connected to a Typewriter Keyboard with 1-Line Display (TDD) for 1-on-1 or ASCII or TDD Modem Communication
InfoTouch	Enabling Technologies Company	$4000-$4900	None	Braille or Typewriter Keyboard Connected to a Romeo Brailler and a Typewriter Keyboard with 1-Line Display (Superprint TDD) for 1-on-1 or ASCII or TDD Modem Communication
TeleBraille	Telesensory Corp.	$5500	None	Braille Keyboard and 20-Cell Paperless Braille Display Connected to a Typewriter Keyboard with 1-Line Display (Superphone TDD) for 1-on-1 or ASCII or TDD Modem Communication

BEST COPY AVAILABLE

page with 6 dots per cell could allocate just under 0.08 watts per dot; 8 dots per cell lowers that to 0.06 watts per dot. An 80-cell by 25-line Braille display, which could provide full text access to an IBM-compatible personal computer screen, could allocate just under 0.04 and 0.03 watts per dot, for 6- and 8-dot cells, respectively. Without any blank lines, a page of Braille text could be expected to have an average of 2 dots raised per cell, so, if only raised dots require power, that would mean the typical power available per dot would be about 3 times the minimum values for a 6-dot cell (4 times the minimum values given for an 8-dot cell). Unless the display is being used for graphics, it would be unrealistic to expect all dots to be raised at once. On the other hand, it would be unwise to design a display so that raising all the dots would blow a fuse or trip a circuit breaker in the user's home or office. A compromise may be necessary for very large displays but it is desirable to have the capability to raise or lower all dots simultaneously.

Applying continuous power to the actuators is impractical for many Braille display actuator technologies because even a one-line display would require more power than a wall socket can provide; far more than a portable battery system could tolerate. Therefore, many paperless Braille displays raise or lower dots and then lock them into position until another page is displayed. Historically, the locking and unlocking mechanisms have required little or no power except while displaying a new page. In practical operation, these locking mechanisms reduce average power consumption by several orders of magnitude. The problem with locking mechanisms has been that they increase mechanical complexity, which tightens the manufacturing tolerances. Reliable actuators for Braille cells are available today but most of them require a locking mechanism to avoid excessive power requirements. Even if power constraints could be ignored, some actuators' locking mechanisms double as a way of ensuring that dots are raised to a uniform height, which is a requirement for Braille.

Designs that employ locking mechanisms update the display dot by dot, cell by cell or in small groups of cells. This minimizes the peak power consumption by decreasing the display update rate. Alternatively, a storage device could slowly accumulate energy from the power source and release it all at once; which is how a portable camera flash works. With the more energy-intensive actuator technologies, a tradeoff is necessary between display size and refresh rate of the display. The inherent size and weight of most Braille display technologies usually justifies slowing the display update time moderately. Portable Braille displays are almost certain to require tradeoffs in power vs refresh rate because both average and peak power capabilities of batteries are strictly limited by acceptable battery size, weight, and frequency of replacement or recharge.

According to a 1990 Smith-Kettlewell study, a one-line Braille display that slides up and down a "page" provides some of the advantages of a full-page display. The study, by TiNi Alloys, Oakland, California, and Smith-Kettlewell, San Francisco, suggests that a six inch long virtual page can even create the illusion of a full size page of Braille. This work may lead to an alternative approach to providing the feel of a full page Braille device in a simpler and more reliable format.

Solenoid electromagnetic actuator technology has been most often tried for producing Braille. Tight packing is needed for displays of useful size, even with coil assemblies and components fabricated with truly miniature solenoids. Historically, solenoids have been power-intensive (requiring locking mechanisms) and prone to failure with dirt from normal use (i.e., grease, skin cells, pollen, and even volcanic ash). This leads to reliability problems because cleaning 6000 solenoids regularly for a full page display would not be a realistic option. Covering the solenoids with a protective plastic membrane keeps the solenoids clean, but slightly moist fingers skip across plastic so a plastic surface is undesirable. Power requirements and interference between neighboring solenoids are also problems that must be overcome. Developments in superconducting materials, and in motor and solenoid miniaturization, may help to solve the problems associated with large electromagnetic Braille display fabrication.

Metec (Stuttgart, Germany), EHG (Nordstetten, Germany), and Tiflotel (Calolziocorte, Italy) have each produced electromagnetic (solenoid) Braille cells that are scalable to multi-line or full-page displays. But the technology was less than successful because of a combination of reliability and repair problems. Power requirements may have also been a factor. Clarke and Smith International (Surrey, England), has produced small quantities of electromagnetic Braille cells, but they were limited to two-line displays. Novanik (Karlstad, Sweden) was working on a 42-cell 29-line electromagnetic display as of 1987, but its status is unknown. Generally, companies seem to have given up on using electromagnetic actuators for Braille. Smith-Kettlewell's proprietary design, described later in this document, is the exception.

Piezoelectric benders, sold in the U.S. are used in all mass-produced refreshable Braille displays with more than a few characters. Called bimorphs, the benders can be made with any of several materials. Lead zirocnate and lead titanate ceramics seem to be the most popular for Braille cells but other piezoelectric materials include single crystals such as Rochelle salt and ceramics such as barium titanate. Piezoelectric materials flex in the presence of an electromotive force. The piezoelectric Braille cells made by Telesensory in Mountain View, CA, are considered by many visually impaired people to have the best feel of any Braille cell available in the U.S. The Tieman cell, another popular piezoelectric Braille cell, is probably made by Kogyosha (Tokyo) working with Braille Equipment Europe in the Netherlands. It is used in the Alva and Braillex displays, and possibly the KeyBraille and Notex displays. Metex, and possibly other electromagnetic Braille cell manufacturers, have switched to making piezoelectric Braille cells. The current state of the art in piezoelectric benders limits refreshable displays with horizontal benders to one or two lines, and only single-line displays are commercially available. The reason for the size limitation is that piezoelectric benders bend very little per unit length, so they have to be much more than an inch long to obtain the bending motion necessary to lift the Braille dots into place.

Elinfa, in France, developed vertical piezoelectric benders for Braille cells, reducing the area required per dot by a factor of five. In theory, the Elinfa cell used in the Personal Touch, could be used to produce a very large display but, to date, producing Braille cells with horizontal or vertical piezoelectric benders has been expensive and labor-intensive.

New manufacturing technologies may be needed to overcome this problem. Ceramic piezoelectrics tend to be brittle, and single-crystal piezoelectrics tend to have other undesirable properties. Single-crystal Rochell salt, for example, has the strongest piezoelectric effect known; but its dielectric properties have led to the use of ceramic piezoelectrics instead. The retail price of piezoelectric displays is $20-25 per dot, which would easily put the retail price of a full-page piezoelectric Braille display over $100,000 per unit. This is far outside the price range of the typical user.

Five factors have enabled piezoelectric displays to dominate the paperless Braille display market. First, although driving piezoelectric cells requires on the order of two hundred volts direct current (VDC), the average current is low enough that they have *very low net power requirements.* The Telesensory Navigator's power supply would only allow a maximum of 0.023 watts per dot. Power consumption is so low that all dots can be raised continuously, thus eliminating the complexity and reliability problems associated with locking mechanisms. Piezoelectrics are actually their own locking mechanisms, requiring no power to stay in position except to cancel leakage currents. Also, low power consumption allows the option of portable battery power for small displays. Second, each dot has few moving parts, the bender and the dot shaft, in the case of telesensory's Braille cells and there is no friction-based locking mechanism. The result is that piezoelectric displays are relatively immune from dirt and wear, though dirt can cause dots to stick, requiring ultrasonic cleaning. Piezoelectric displays are reliable due to *a minimum number of moving parts and minimal friction*. Third, piezoelectric displays can provide *fast display updates* because they are energy-efficient. Fourth, piezoelectric displays are *very quiet*. They make just enough noise to let the user know an update has occurred. Finally, the dots can be closely packed and therefore come very *close to the standard Braille dimensions*.

The next generation of full page Braille displays must be able to provide refreshable Braille for significantly less than $20-25 per dot. It is very unlikely that a full-page Braille display could be sold for much more than the $15,000 price of existing 80-cell displays. This means a 25-line, 40-character display with 6 dots per cell, would have to be produced at a cost of $2.50 per dot. A significantly lower price, perhaps $1 per dot, would open up a much larger market for full-page Braille displays and serve many more persons with vision impairments.

Tactiles is working on a machine with very low cost self-locking dots (<$.10) and a travelling "printhead" similar to a dot matrix printer. The target price is under $4000.

The reliability needed for every dot in a Braille display is significant. For example, if every dot in a full-line Braille display (480 dots) worked 99% of the time the display would be error-free once in 125 displays. If the dots were 99.99% reliable, the 80-cell display would be error-free only 95% of the time and the full-page display would be error-free about 55% of the time. At 99.9999% reliability, the one-line display would have an error every 2000 lines, but the full-page display would still have an error once in about 165 pages. This makes the design of a reliable full-page display difficult. Moving parts tend to make a device unreliable, but a Braille display must have 6000 independently moving parts, each with reliability much greater than 99.99%. Telesensory's Braille cells have been around for a long time and are extremely reliable, but there have been considerable differences in reliability among the manufacturers. Also, characteristics of the user such as sweaty hands and a tendency to eat potato chips, affect reliability.

A major technology shift is required to design a full page Braille display to meet the media access needs of persons with vision impairments. This new technology shift would incorporate advanced materials and computer control technologies. Advanced materials and manufacturing technology may make it possible to implement several lines of Braille; perhaps a full 40-character by 25-line page of Braille output.

An example of a technology improvement that could facilitate the implementation of full-page paperless Braille is large array controllers for liquid crystal displays (LCDs). These LCD controllers can control 64 high-voltage lines, on the order of 120-180 volts direct current (VDC) from a single chip and could be used to control 10 piezoelectric Braille cells. They may also be useful for switching electrorheological fluids, which will be described later. Other LCD controllers are available that could control 20 or more elements.

Before discussing the newer technologies, a review of earlier attempts at a full-page paperless Braille display is useful to prevent repeating mistakes.

Historical Developments

Thermostat metals were used in Braille Inc.'s prototype Rose Braille Display Reader, a full-page display patented in 1981 but was never commercially produced. The Texas Instruments thermostat metals used are bimetallic strips that bend when heated. In the Rose Reader, the shaft of each Braille dot has a grooved ring around it. The shaft would be pushed up by a spring, but a hook on the end of a bimetallic strip catches the groove on the ring and restrains the dot. When heat bends the bimetallic strip away from the ring, the spring raises the dot. A separate manually activated mechanism pushes all of the dots down again and signals the machine to display the next page. The unit included a panel of 12 control buttons and a cassette drive for storing text.

According to Leonard Rose, the one prototype that was built had some dots that did not work because the device was handmade; possibly machine-made parts would have been more reliable. Unfortunately the only accurate way to measure reliability very near 100% would be to manufacture parts. This would require a considerable investment. According to Mr. Rose, putting the system into production would cost about $750,000 with units eventually selling for as little as $7500. The thermostat metals used in the Rose Reader are less expensive than piezoelectric elements, and can be designed in modular units for easier repair. However, the number of moving parts per dot, the direct use of heat and friction, and the use of a manual mechanical reset mechanism are all potential sources of reliability problems. In principle, replacing the manual reset lever with an electric one is easy, but that is likely to affect reliability.

The Rose Reader requires raising the temperature of the metal strips by 30 degrees Fahrenheit to overcome friction. This requires significant power, so the dots are raised one at a time, at a rate of 200 dots per second. If all dots are raised, the total time required for a full display is about 30 seconds, although the average Braille page would take approximately 10 seconds. Based on an estimate from the patent information, if a page could be displayed within one second, it would trip a 15 ampere circuit breaker, even with only one third of the dots raised. Each dot requires on the order of a watt-second of energy to be actuated. This is because the temperature changes required in thermostat metal actuators that have to overcome friction make them relatively energy-intensive.

It is not clear whether sufficiently reliable mechanical locking mechanisms are available to circumvent higher energy requirements, but there seems to be a pattern. Energy-intensive actuators with lower materials costs tend to require locking mechanisms that cost as much to manufacture as more energy-efficient actuators for a given level of reliability. In the end, the prototypes with locking mechanisms are generally too costly to manufacture or too unreliable to sell. When the development funding for these devices runs out, the designs are shelved indefinitely. What is needed is an actuator that is energy-efficient enough to be used without a locking mechanism, yet costs less than $2.50 per dot or less and fits in a standard Braille cell profile.

From the late 70's to the mid-80's, the American Foundation for the Blind (AFB) experimented with injection-molded arrays of 64 x 64 dots, manipulated one row at a time by a single row of 64 solenoids, one row at a time. Four of these prototypes were built with the combined capability of producing a full page of Braille or graphics. Graphics capability turns out to be a mixed blessing because evenly-spaced dots are incompatible with standard Braille dimensions but would be a major advantage of a full-page display over a one-line display. The system had three major advantages: the pins were mechanically latched into position so power consumption was moderately low (because the system was slow); the feel of the display was good; and since the system was modular, a mechanism for repair by replacement was provided. There were two major problems: the 64-step display update was slow enough to offer no great advantage over paper output, and the system was expensive. Ultimately, the cost of the mechanical system was its downfall, and the method was not recommended for further development.

Shape memory alloys were the technology used by TiNi Alloys, to develop a 20-cell by 3-line prototype display of 8-dot cells. Shape memory alloys are nickel-titanium alloys that forcefully return to a preset shape when heated and are usually alloys of nickel and tin, or of copper, zinc and aluminum. In this case, TiNi used a nickel-titanium (nitinol) allow in the form of a wire, one inch long and 0.030 inches in diameter. When an electric current heats the wire, it shortens, pulling the shaft of a dot down against the force of a spring. Each dot has a small flexible piece of sheet metal with a hole big enough to let the shaft of the dot pass through it if the metal is lying flat, but small enough that it catches the dot if the piece of metal is angled. The metal's resting position is angled. When a dot is lowered, its shaft catches on the piece of sheet metal and pulls it down flat enough to let the shaft pass through. When the wire cools, it stops pulling the shaft of the dot down, and the dot tries to spring up again, but that pushes the sheet metal back to its angled position, catching the dot's shaft so it cannot move. When a plate pushes all of the pieces of sheet metal flat, the dots are then free to move and all raise, clearing the module. Stops were used to give the display adequate feel. The display was built of 4-cell units because modularity makes it easier to build and repair a display. Funded by the National Institute of Health, the project did not get past the prototype stage, though it received an Excellence in Design award from *Design News*, a respected journal for design engineers.

The technology was capable of displaying a full Perkins Brailler page. In fact, the software was written for an 80-cell by 25-line display but there were reliability problems with the detent and release mechanisms, which required tight manufacturing tolerances. With further development, the technology might have become cost-competitive with other Braille cells, but the two-year grant ended in 1990. An important cost driver was making and attaching the special metal wire, although better techniques are now available. Power requirements were 50 watts instantaneous, or about 1 watt per dot. The dots were given that power, one module (32 dots) at a time, for 50 or 60 thousandths of a second, so the energy requirement per dot is 0.050 to 0.095 watt-seconds. That is much better than thermostat metals, but not as efficient as piezoelectric technology. Shape metal alloys were a valuable experiment for paperless Braille, but their cost and power performance seem unlikely to do much more than match piezoelectric technology.

Current Developments

For several years, Smith-Kettlewell has been working on a proprietary electromagnetic Braille cell technology funded by NIDRR. It is limited to displays of 80 characters or less, but the cost is estimated to be $20 per cell, which is below the cost of piezoelectric displays. Few details are available, but the technology has a fast refresh rate and has the potential to be used in portable systems operated from battery power. Smith-Kettlewell has passed the design on to a developer, though no estimated date for production was given.

During the past six months to a year, Blazie Engineering, Street, Maryland, has been working on a pneumatic display that uses puffs of air to move tiny bearings supporting Braille dots. No product is anticipated until 1993. The device requires a spacing approximately 0.015 inches greater than the standard distance between Braille cells. The display is supposedly scalable to a full page. Preliminary cost estimates are as low as $5 per cell, which would be less than $1 per dot. The feel of the display is said to be solid, and it is expected that the present refresh rate can be increased. Power requirements are predicted to be low, and a 20-cell prototype has been built. The display has two moving parts per dot and can be cleaned by immersing it in liquid. Performance and cost predictions based on the prototype must be considered preliminary.

Recent advances in sequential soft-copy Braille displays have been made by Tactilics, Inc. and Densitron Corporation. Sequential soft-copy Braille displays are essentially belts that move across a "window" while Braille dots are raised on their surface. Densitron has been selling prototypes of a 40-cell deformable plastic disposable belt device for $2995. Its lack of navigability would appear to limit its use.

Tactilics' belt is made of hard, molded nylon cell sections which they indicate is long lasting and self-cleaning. They claim its bi-directional control makes it highly navigable and that it is a true realtime "monitor." Also, when battery powered, the unit may be used as a portable "book" and that its mixture of high and low tech is a price breakthrough. Two units will be introduced in mid-1992: 1 50-cell for $1200 and an 85-cell model for $1500.

Future Developments

What lies beyond the existing systems is impossible to predict with certainty because, though completely new technologies are seldom discovered, old ones are constantly revitalized by new computer capabilities, materials, and manufacturing processes. Sometimes older technologies suddenly become practical due to material or other technology breakthroughs. Some companies are unwilling to discuss technologies they are considering for paperless Braille. Blazie Engineering suggested three: magnetostriction, electrorheological (ER) fluids, and polymer gels.

Magnetostriction is the property of some alloys that cause them to forcefully expand in a strong magnetic field. "Giant" magnetostriction, an expansion on the order of 0.15%, occurs in alloys of certain rare-earth elements. An alloy of iron with terbium and dysprosium is used in an actuator sold by Edge Technologies in Ames, Iowa. There are serious problems with using that technology for Braille. Rare-earth elements are not really rare, but they are expensive to purify. The alloys used must be in the form of a single crystal which is presently expensive to refine and produce. Finally, the effect of magnetostriction is too small to be used directly without long pieces of the alloy, which may be both voluminous and cost-prohibitive.

Levers are being tried for converting some of the force from the actuators to linear displacements that would be adequate for Braille. A hard limit seems to be that the cost of an individual actuator is still not competitive with piezoelectric technology, and the cost of coils to produce the magnetic field, exceeds the cost of the special alloys as the actuators gets smaller. As with piezoelectrics, benders can theoretically replace levers as a way of trading force for increased movement, but the single crystals are brittle. No one knows if their tensile strength is such that the elements will break if used in benders. Power requirements are estimated at a maximum of 10 watts per dot, which is high, but locking mechanisms may allow power management strategies with low average power requirements.

Other magnetostrictive materials exist, but it is not clear that any provide enough of an effect to be useful for Braille cells. Magnetostrictive ribbons, which are used in sensors, provide extremely high efficiency, but they lose most of that efficiency in strong magnetic fields, thus limiting their maximum expansion. Magnetostriction has its highest efficiency when the actuator is moving back and forth rapidly, though that problem might be solvable by making the dots move back down slowly, thus extracting a displacement as close to the maximum as possible. In summary, magnetostriction does not appear to be a cost-competitive technology for Braille cells, though this may change with future breakthroughs in materials technology.

Fundamental research on electrorheological (ER) fluids is being conducted at the University of Michigan (UM), Ann Arbor. In 1988, a UM scientist made an important breakthrough in electrorheological fluids development. ER fluids thicken when a strong electric field is applied to them, on the order of 2000 volts per millimeter. Their consistency changes from liquid to something "more like Velveeta cheese." This allows hydraulic actuators to be constructed. ER fluids stop flowing while in a strong electric field, so, as a hydraulic fluid, they can selectively apply pressure to actuators. Per hydraulic switch, power requirements are lower than piezoelectric technology but a pump is required to supply the pressure for the hydraulics, and therefore their overall efficiency is unclear.

The breakthrough at UM was to find an inexpensive ER fluid that does not contain water. Water content lowered the efficiency and predictability of previous ER fluids making them impractical for actuators. The fluids used at UM are inexpensive but the particles suspended in them tend to separate from the liquid. More expensive ER fluids do not have this problem. Three problems are likely to arise with ER fluid-based Braille cells: the use of a liquid, difficulty with modularizing a system with fluid lines, and fluid pump power and noise. Without modules, a large Braille display could be very difficult to build and repair. There may be ways to modularize a hydraulic display. Pump power and noise may not turn out to be an issue, but the use of a liquid seems likely to be a challenge to developers. The need for intense electric fields could be reduced by using narrow gaps. Overall, ER fluids may be feasible for Braille cell development in the immediate future.

Polymer gels are another promising technology for full page Braille displays is polymer gels. Polymer gels collapse when exposed to intense light. These gels are being developed at the Massachusetts Institute of Technology's (MIT) Department of Physics and Center for Materials Science and Engineering. Under the proper conditions, gels can be induced to reversibly release a large portion of their liquid content. This is called collapsing, because releasable liquid content increases the volume of a gel by factors ranging up to 350 or more and multiplying their length, width and height by a factor of 7. In 1990, researchers at MIT induced a light-absorbing gel to collapse by heating it with a visible laser after having induced gels to collapse with exposure to ultraviolet rays, voltages on the order of 5 volts, and changes in the surrounding liquid's temperature, composition, pH, and salt content. The visible light has the advantage of safety and speed over ultraviolet radiation, as well as providing a controlled way to induce small temperature changes through a sealed container. The sealed container is necessary because, to reabsorb the liquid, a collapsed gel must be immersed in the liquid. The liquid and gel have to be separable to exploit the volume change of the gel, but gel reaction times below a second require gels significantly thinner than a human hair.

To manage fine fibers, researchers in Japan have formed gels into sponges or bundles of fibers, but reaction times are still greater than one second. The leading light-sensitive gel researcher at MIT, Dr. Toyoichi Tanaka, estimates that strands of gel one thousandth of a millimeter in diameter, about the diameter of muscle fibers, would react as fast as muscle tissue. Doubling the diameter of a gel quadruples its reaction time, though; some gels react very slowly with modest increases in fiber diameter. So far, heating a gel with a laser is moderately power-intensive.

Minimal research and development could conceivably reduce the power required by several orders of magnitude. The choice of gel material, the concentration and choice of light-absorbing material in the gel, and other factors could significantly reduce power requirements. According to Tanaka, red lasers work better than the violet-blue laser that was used to estimate power efficiency. Diode lasers with power efficiencies of over 30% are now available that emit red light. Though diode lasers with lenses cost on the order of $50 a set, and that is for lower power and in quantities of a thousand, if one or more lasers were scanned over a Braille dot array with a mirror, gel technology might make it possible to implement a reliable full-page Braille display with reasonable cost, size, weight, and power.

Tight temperature control would be a potential problem (i.e., temperatures held within ± one degree), but the MIT researchers have experimented with gels at room temperature without temperature regulation, and in water. The MIT group uses a low-cost gel, which is also encouraging. The temperature at which a gel collapses can be controlled by the proportions of two liquids into which the gel is immersed. Even if some temperature regulation turned out to be necessary, advances in solid-state Peltier effect heating/cooling might be applicable, balancing the power requirements of laser(s) with those of a temperature regulation system. It is too early to predict whether the feel of a gel-based Braille cell would be adequate, but the technology shows promise with additional research and development.

According to "Tactile Displays for the Visually Disabled--A Market Study, July 1987," published by the Swedish Institute for the Handicapped, materials that expand with moderate heating have been tested for application to Braille displays, apparently without success. That reference does not indicate what material was tested, but it would not have been a polymer gel. Gels could be used for a phase change, but that phase change could not be accurately described as a transition from a solid to a fluid.

The piezoelectric materials currently used in Braille cells are not the only ones available. A study of recent alternatives would be worthwhile. For instance, A.V.X., in Myrtle Beach, South Carolina, started selling lead zirconate titanate (PZT) in multiple layers around the beginning of 1991. It appears to be possible to get actuators made from this material for less than $20 each, in quantity, making multilayer PZT marginally competitive with existing piezoelectric displays. Layering reduces voltage requirements, but it is unclear whether the increased materials costs would be justified. A tough plastic film, called polyvinylidene difluoride (PVDF), sold by Atochemwith the trade name of Kynar, may also be useful for Braille cells. It can be used at very high voltages, compared to ceramic piezoelectrics, but its efficiency is lower than that of ceramics. It is apparently better suited to vibrating Braille dots than static ones, for reasons that include power efficiency, but the feel of vibrating Braille dots is not as good as the feel of static dots. Further study is needed to determine whether these and other materials are appropriate for use in Braille cells. In particular, their response at low frequencies must be taken into account.

Superconducting magnets will eventually facilitate the miniaturization of solenoids because strong superconducting magnets can be fabricated in a small package without the overheating, even when densely packed. By definition, electricity running through superconducting materials produces no heat, which makes superconducting magnets incredibly energy-efficient. So far, the "high temperatures" required to use high-temperature superconductors are on the order of 300 degrees below zero Fahrenheit, but they are slightly higher than the temperature at which nitrogen, the principal component of air, liquefies. Liquid nitrogen, which is used to keep existing high-temperature superconductors cold, has been touted as being cheaper than beer, but anything that cold must be treated with extreme caution in devices developed for the general public. Liability issues could be enough to eliminate superconductive solenoids from consideration for Braille displays. Also, the known high-temperature superconductors are brittle and somewhat expensive, so even the discovery of superconductivity near room temperature, if that is possible, would not guarantee applicability to Braille cells. Superconducting materials applications are in the fundamental research stage and it will be 5 to 10 years before applications are marketed.

An entirely different approach to paperless Braille uses electrodes to indicate the presence of a Braille dot with a tiny electric shock below the threshold of pain. This approach has potential for low cost, high speed, and small size, but experiments have not produced a design acceptable to the end users. Until technology is perfected, it cannot be considered a viable technology for Braille displays. Problems include reduced reading speed and very wide variations in skin resistance, both among different people and with sweat and other factors.

An alternative approach to high reliability would be to use what are called "smart" materials. Smart materials combine sensors and actuators to react to special situations. In this case, smart materials might be able to provide high reliability with imperfect locking mechanisms by verifying that a dot has been raised or lowered. If not, an actuator could be triggered several times, allowing the reliability per trigger to be lower. Alternatively, the actuator could be vibrated to free any dirt that caused the failure. The smart materials approach is a compromise, attempting to avoid the high cost of piezoelectric technology which does not need a locking mechanism against the high cost of an extremely reliable locking mechanism. The smart materials approach would give a reliability boost to a mechanism that is already reliable. It would not be adequate with a mechanism that has a high failure rate. This is because if a dot fails to work after perhaps two or three tries, then the display's electronics would have to indicate an error. When a display operates properly, the error message should not appear except in the case of a catastrophic failure. The smart materials approach might also be used to adjust the overall height of the dots on the display. The cost of this feature may be prohibitive. At this time, shape memory alloys would probably be the best choice for testing along with low-cost piezoelectric-film sensors.

Telesensory's Optacon II bears mentioning because the piezoelectric device provides access to printed text and the technology might be adapted to Braille in some way in the future. Based on ceramic piezoelectric technology, it uses 100 vibrating rods (5 columns by 20 rows) to present the image of letters from a small camera. Many persons with visual impairments find the Optacon II useful for reading print and interpreting graphics. It provides instantaneous results and offers great flexibility and portability. Like Braille, learning to use the Optacon II takes many hours of training and not everyone can become proficient in its use. The Optacon II is intended to supplement Braille, not replace it. Embossed Braille requires no special reading equipment or equipment costs, and it was also found to be easier and faster to read than raised letters. The Optacon produces raised letters when used to read print.

It is important to note that the skin's sensitivity to fixed dots differs from its sensitivity to vibrating dots. Vibrating dots are not used for Braille because vibration temporarily reduces the skin's tactile sensitivity and actually reduces the ability to read tactile information. This is unfortunate because many piezoelectrics move at least an order of magnitude more efficiently near their resonant frequency. Vibrating dots also generate a buzzing sound, but that is a solvable secondary issue compared to the human factors problem.

A detailed discussion of tactile graphics displays is beyond the scope of this document, but they are closely related to Braille displays. The big difference is that they generally use an array of evenly-spaced dots instead of the standard Braille spacing. These displays offer the advantage of graphics capability, but, in general, they cannot produce Braille with standard spacing.

This document has concentrated on paperless Braille technology. However embossed Braille technology is also important. The primary technology for embossing Braille is solenoids, and this seems likely to continue for many years. The only commercial alternative has been the use of molten plastic containing magnetic material. The plastic is magnetically guided onto paper to form Braille dots. Until 1990, Howtek sold an embosser, the Pixelmaster, based on this principle, but it could only raise dots eight thousandths of an inch. Even twelve thousandths of an inch is barely tolerable for Braille, and far below standard dot height. So the Pixelmaster, which could also produce printed text, was considered appropriate only for tactile graphics and visible print. Technologically, the Pixelmaster could have probably produced Braille of normal height, but it was originally designed for advertising. The plastic "ink" was originally intended for producing more brilliant colors displays and not to produce Braille. Similar technology was tested in Japan and found to produce the proper height of dots, although dot shape was a problem. Plastic dots on paper are much more durable than conventional embossed paper dots, but it is unclear whether their feel can be made acceptable to the users. Dots that are too smooth are harder to read.

Smith-Kettlewell is in the very early stages of developing a thermal embossing technology, said to offer the possibility of fast and silent paper Braille.

Multiple copies of Braille text can be made with heated plastic sheets that conform to the shape of the original Braille page. This process is called vacuum forming. In larger quantities, printing press techniques can be used to emboss Braille, but copying Braille is still expensive. Even in quantities of a few hundred, a relatively fast Braille embosser still has advantages over Braille copying technology, including the feel of paper vs plastic.

Sheets containing encapsulated ammonia are being used to produce some tactile graphics, but the special plastic sheets required are too expensive to be used as a substitute for paper Braille. They are used with thermal copiers, but have the feel problems associated with plastic too.

10.0 COST CONSIDERATIONS OF ADVANCED BRAILLE TECHNOLOGY

As explained in the preceding section, advanced technology Braille must cost less than 20 to 25 dollars per dot to be cost effective when compared to existing technology. A substantial price reduction would be needed to make multiple-line displays marketable, and a 40-cell by 25-line full-page display would have to cost less than $2.50 per dot to be in the price range of existing market forces.

Piezoelectric displays, the dominant technology at this time, show little hope of dramatically dropping in price in the near future unless a much less expensive material for them is found or new manufacturing techniques formulated. Thermostat metals, though cheaper, are not energy-efficient, making them a poor choice for displays. Highly mechanical approaches tend to be expensive, unreliable, or both. Shape memory alloys tend to be somewhat expensive, though moderately energy-efficient.

Smith-Kettlewell's electromagnetic technology seems to be able to offer a significant price breakthrough, but it is limited to one-line displays.

Soft Braille offers an apparently inexpensive alternative to refreshable Braille technology, but its costs may be misleading. Refreshable Braille makes it feasible to move to different points in a document without too much confusion, but Soft Braille is best suited to cover-to-cover reading which is not the reading pattern except for some pleasure reading. Also a disposable belt model must be judged on the basis of the cost and inconvenience of replacing the belt.

On the horizon, magnetostriction seems to offer little hope for a price breakthrough at this time. On the other hand, ER fluids seem to have the potential of producing a low cost Braille display. However, there are significant engineering hurdles to be overcome.

Polymer gels offer future hope of low-cost displays, though they are still in the basic research stage. New developments in piezoelectric materials merit further study. Superconducting solenoids and silicon micromachines are probably beyond the ten-year scope of this study, and it is not clear whether electrode-based Braille will ever have the feel demanded by persons with vision impairments.

Based on the outlook for affordable full-page Braille displays, two compromise approaches should be considered. First, smart materials are a way of increasing the reliability of existing mechanical locking mechanisms: Eliminating the need for locking mechanisms might be better in the long run, but that may not be technologically feasible without the expense of piezoelectric displays. Second, further investigation into the effectiveness of a sliding one-line display is justified by the lack of compelling evidence that a full-page Braille display is technologically feasible in the next three to five years. However, research and development efforts should be devised to push the technologies described above beyond the laboratory.

11.0 COST BENEFITS TO PERSONS WITH SENSORY IMPAIRMENTS WITH EARLY INCLUSION OF BRAILLE

Braille displays are presently costly, and therefore many persons with vision impairments currently use synthesized speech instead. More affordable Braille displays would give persons with vision impairments more options between speech and Braille, and increase literacy among persons with vision impairments. Lower-cost Braille displays would allow larger displays to be purchased per dollar. Perhaps the biggest short-run cost benefit would be the improved earning potential that a better Braille display could give blind workers, especially in the computer programming and office environments.

In the long run, an affordable full-page Braille display would contribute to Braille literacy, education of the blind, and access to computers, empowering persons with visual impairments.

12.0 PRESENT GOVERNMENT INVOLVEMENT IN ADVANCED BRAILLE TECHNOLOGY

At present, no Government programs are known to be supporting Braille display development, although virtually all previous development have been Government-sponsored.

13.0 ADVANCED BRAILLE TECHNOLOGY TIMELINE

Paperless Braille technology has settled on piezoelectric actuators because of their reliability and energy efficiency. To compete with piezoelectrics, any new technology must provide reliability and energy efficiency at a lower cost per cell. As computers become increasingly portable and dependent on battery power, only the most energy-efficient Braille displays can be used in these portable units. As the most energy-efficient proven technology for Braille displays, piezoelectrics should be reevaluated to determine if new manufacturing technologies can lower the cost of piezoelectric displays.

Smith-Kettlewell's new electromagnetic technology, already in the process of becoming a commercial product, may provide significantly less expensive one line displays. Expected in 1993 is Blazie Engineering's pneumatic display which may make full-page displays affordable.

Recent Soft Braille displays offer a compromise between low cost and limited performance. Their price is revolutionary, but their capabilities are extremely limited for most applications other than reading for pleasure.

Magnetostriction is not likely to be a major factor in Braille displays, but clever designs or new materials breakthroughs could overcome the cost barrier. ER fluids, with the potential for low cost and high energy efficiency, should be ready to begin development in 1992. Polymer gels should be available in actuators between 1993 and 1994. If their energy-efficiency proves to be high enough to make them practical, they should be evaluated for Braille displays. However, they are still in the basic research phase in 1991. Superconducting solenoids and silicon micromachines are beyond the ten-year time scale being considered by this study, and electrode-based Braille seems unlikely to be practical at this time.

A smart materials approach, with shape memory alloys and piezoelectric film sensors, merits some consideration, but it offers only a limited chance of success in competing with piezoelectric displays. That approach is moderate-risk, moderate-gain, and probably moderate-cost, since some of the development has already been done.

Experimenting with a one-line display that slides up and down a page, that may be compressed vertically, is a higher-risk approach, but it offers potentially very high gains if it makes a full-page display unnecessary. Cost should be relatively low in this approach, making it especially attractive.

14.0 PROPOSED ROAD MAP FOR INCLUSION OF BRAILLE CAPABILITIES

The U.S National Aeronautics and Space Administration (NASA) and the Department of Defense (DOD) fund actuators for specific systems. Neither is in the business of developing advanced actuators for Braille. However, in a cooperative funding teaming arrangement for research and development on small lightweight actuators there probably would be a great deal of interest. For example, both the DOD and NASA fund research and development efforts for devices for the handicapped through small business and university innovative research grant programs.

The Department of Education, as a first step in Braille cell development, should explore the possibility of a cooperative effort with both NASA and the DOD in the area of small actuators for use in full-page Braille cells. This would lead to a research and development program for small low power consumption actuators for use in braille display devices, robotics, and space applications.

A program would then be established to develop single-cell actuators that could be used in braille displays. NASA or the DOD could sponsor the initial basic research under small business or university educational grants for cells of one to six elements. This could be followed up with a program by the Department of Education under a grant for a Phase I program to develop a single-line Braille display of 80 characters. Finally, a grant would be awarded for an 80 column by 25 line display.

Key to each phase would be three to four research organizations competing for the next phase of the program. For example, NASA might start with four to five organizations for a one year concept design study at approximately $125,000 per year, for small low power actuators based on advanced technology. Designs would be sought for both single actuator and multiple cell designs. NASA would then select the three organizations with the most feasible designs for a second phase to integrate the individual actuators into a prototype Braille cell display and fabrication effort at $150,000 each for one year.

The basic research effort would then be followed up with a Department of Education effort to integrate one or two of the organizations' actuator designs into low-power, full-page Braille displays over a two to three year program. The Department of Education and NASA would jointly fund the efforts and cooperate in the exploitation of the technology.

Finally, the Department of Education would fund a production study and transition the devices to full scale development.

15.0 POTENTIAL PROGRAM SCHEDULE

Figure 2 presents a potential program schedule for the development of a braille display unit using advanced technologies. The Department of Education could act as the program administrator, with NASA and the DOD providing basic research and development expertise at critical times throughout the development effort. Within five to six years a full-page Braille display could be on the way to full scale production.

	FY92				FY93				FY94				FY95-97				FY97-98			
	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
Establish cooperative agreements (NASA/DOD)		Δ																		
Actuator R&D Feasibility Studies (NASA) -- Full page -- Virtual page				Δ			Δ													
Actuator Developments (NASA/DOED)							Δ					Δ								
System Integration Programs (DOED)												Δ				→				
Transition Technology to Industry																	Δ			Δ

Figure 2. Full Page Braille Development Schedule

INPUT/OUTPUT DEVICES FOR COMPUTER AND ELECTRONIC BOOK ACCESS

MARCH 1992

Prepared by

Daniel E. Hinton, Sr., Principal Investigator
and
Paolo Basso-Luca and Charles Connolly

SCIENCE APPLICATIONS INTERNATIONAL CORPORATION
3701 N. Fairfax Drive, Suite 1001
Arlington, VA 22203
(703) 351-7755

1.0 SCENARIO

Input/output devices for computer and electronic book access.

2.0 CATEGORY OF IMPAIRMENTS

Persons with vision impairments.

3.0 TARGET AUDIENCE

Consumers with Vision Impairments. Persons with vision impairments will benefit from enhanced access to media information services. This scenario on advanced information storage technology for printed material will provide a means to disseminate information to consumers with vision impairments on how electronic books could effect their media access. In particular, it will provide a better understanding of the technology available in electronic media over the next three to five years and the potential problems that could arise in media access.

Policy makers, including national representatives, government department heads, and special interest organizations. Policy makers will also benefit from this scenario because they can apply this scenario to better understand the issues related to electronic media access for persons with vision impairments. In addition, this scenario provides a point of departure for them to understand how advanced technology funding priorities within Government programs can accelerate access for persons with vision impairments to the ever expanding field of electronic media storage and retrieval of information. It will also provide a point of departure for legislation or regulatory action necessary to ensure electronic books and other electronic medias are accessible to persons with vision impairments.

Research and Developers. This group will benefit through a better understanding of the needs of persons with vision impairments and specifically their printed media communications requirements. This understanding of media access requirements will assist researchers and developers in designing media access functions in their future products to meet the needs of persons with vision impairments.

Manufacturers. Manufacturers will benefit through a better understanding of the potential market size and the existing Federal Government requirements for media access for persons with vision impairments which can be met by adding an access capability to their electronic media products.

4.0 THE TECHNOLOGY

Several microcomputer based technologies have impacted the way visually impaired people access information from the computer and electronic books. Visually impaired individuals have their primary difficulties with the output display, although newer display-based input systems (e.g., mice, touchscreens) may also pose problems. This group includes individuals who have failing vision and individuals with partial vision, as well as those who are blind. The primary solution strategies involve providing a mechanism to connect alternate display or display translator devices to the computer, and providing alternatives to display-based input.

Technologies which are associated with computer input devices include braille keyboards and optical character recognition. Blazie Engineering's Braille 'n Speak is a pocket talking computer with speech hardware and software built into the unit, and a braille keyboard. It also acts as an input device for DOS machines. The user presses a combination of keys that produces a standard 6-dot braille symbol. Optical character recognition is covered in a separate scenario titled, "Character Readers for Dynamic LED and LCD Display Access." Table 1 provides a representative list of optical scanners on the market today.

Current technology output devices include: Braille output systems, speech synthesis, and large-print processing. Braille output systems are covered in a separate scenario titled, "Braille Devices and Techniques to Allow Media Access." Synthesized speech is one of the most powerful and least expensive access devices for the blind. Generally, a speech system consists of resident software that converts text into speech, a speech-synthesis board with audio amplification and an interface to the PC bus, and a speaker that sits outside the computer. When users press a series of keys on the keyboard, the system turns the letters into phonemes (the smallest units of sound), runs through a series of rules that tell it how to say the word, and outputs the word through the external speaker. Tables 2-4 show the numerous speech and audio products on the market designed specifically for persons with vision impairments. Also available are many screen reader-software packages designed to direct keyboard input and screen text output directly to the voice synthesizers.

Large-print processing is a valuable access medium for the visually impaired. Individuals with low vision may have difficulty reading the screen because the characters (text), or images are too small. In addition, they may have difficulty seeing the screen due to glare or distance. The two basic methods to add large print to an existing personal computer are to connect a hardware-based large-print processor or load a software package that remaps the characters of the video display to increase the size of the characters of image. Current hardware products are depicted in Table 5. Hardware-based large-print systems use a special video card, a larger monitor to increase font size, and a special joystick or mouse to move the cursor around the screen. The software-based large-print systems provide larger letters and graphics without any additional hardware but impact the speed of program execution. Current software products are depicted in Table 6. These software based systems support a variety of computer functions such as word processing, graphics utility, printer utility, and braille word processing.

Table 1. Optical Character Recognition Devices

Brand Name	Manufacturer	Price	System	Description
PC/KPR	Kurzweil Imaging Systems, Inc.	$3,995-6,995	IBM	OCR system with voice output
Personal Reader (KPR)	Kurzweil Imaging Systems, Inc.	$7,950-11,950	All	OCR system with voice output
Kurzweil 5000, 5100 and 5200 Scanning Systems	Kurzweil Imaging Systems, Inc.	N/A	IBM	OCR system with voice output
Adhoc Reader	Adhoc Reading Systems, Inc.	$6,290	IBM	OCR system with voice output
Arkenstone Reader Models S and E	Arkenstone	$1,495-3,995	IBM	OCR system with voice output
Cannon 1X 12 Scanner, Cannon PC Interface Board and Readright V1.13 Software	Canon USA, Inc.	Scanner $795 Software $595 Board $395	IBM	OCR system
Oscar	TSI/VTEK	$3,895-4,295	IBM	OCR system
Discover 7320 Models 10, 20, 30	Kurzweil Imaging Systems, Inc.	$3,995-6,995	IBM	OCR system
Omni-Reader	IMPX	$199	Apple, IBM	OCR scanner
Totec Model TO-5050 Proscan and TO-5000B	Totec Co. Ltd. Legal Scan Serve, Inc.	$9,990	N/A	OCR scanner
PC Scan 1020 and 2020	Dest Corp.	$1,900-1,945	Apple, IBM	OCR scanner
Deskscan 2000 and 3000	Chinon America, Inc.	N/A	Apple, IBM	OCR scanner
Personal Computer Scanner (PCS)	Compuscan, Inc.	$3,495	IBM	OCR scanner
Scan 300/S	Abaton Technology Corp.	$1,595	Apple, IBM	OCR scanner
Readstar II Plus	Inovatic	$995	IBM, Apple	OCR software
Readright 2.0	OCR Systems	$495	IBM	OCR software
Docread I, III and Expert	Adhoc Reading Systems, Inc.	$2,690-6,290	IBM	OCR software for Adhoc Reader
Read-It	Olduvai Corp.	$295-595	Apple	OCR software

Table 2. Speech Synthesizers

Brand Name	Manufacturer	Price	System	Description
Doubletalk	RC Systems, Inc.	$249.95	Apple, IBM	
Apollo	Dolphin Systems	$687	IBM	
Readme System; Termivox; Termiscreen Reader	Infonox	$1695 $1995 $445	N/A	
Echo+ Speech Synth	Street Electronics Corp.	$119.95-179.95	Apple	
Votrax, Personal Speech System	Votrax	$449.95	Apple, IBM	Voice output module
Accent-MC	Aicom Corp.	N/A	IBM	
Accent-XE	Aicom Corp.	N/A	Toshiba	
Synphonix 230 and 235	Artic Technologies	$595-1,095	Toshiba	
Synphonix 310 and 315	Artic Techologies	$695-1,095	IBM	
Synphonix 250 and 255	Artic Technologies	$695-1,195	Toshiba	
Echo Commander	Street Electronics Corp.	$164.19	Apple	
Synphonix 220 and 225	Artic Technologies	$495-995	Toshiba	
DECTALK	Digital Equipment Corp.	$4,498	All	
Intex-Talker	Intex Micro Systems Corp.	$345	All	Voice output module
Echo II	Street Electronics Corp.	$116.95	Apple	
Artic Crystal	Artic Technologies	$1,195-2,095	IBM	
Audapter Speech System	Personal Data Systems, Inc.	$1,095	All	
Blackboard	Peripheral Technologies, Inc.	$595	Apple	
Calltext 5000	Centigram	$3,225	IBM	
Calltext 5050	Speech Plus, Inc.	$3,900	All	
Echo 1000	Street Electronics Corp.	$134.95	Tandy	
Echo IIC	Street Electronics Corp.	$134.95	Apple	

Table 2. Speech Synthesizers (Continued)

Brand Name	Manufacturer	Price	System	Description
Echo MC	Street Electronics Corp.	$179.95	IBM	
Echo PC+	Street Electronics Corp.	$161.95	IBM	
Personal Speech System	Votrax, Inc.	$449	All	
Syntha-Voice Model I	Syntha Voice Computers, Inc.	$695	IBM	
Speaqualizer	American Printing House for the Blind	$809.41	IBM	
Speech Adapter for PC Convertible	IBM Corp.	$620	IBM	
Speech Thing	Covox, Inc.	$79.95	IBM	
Synphonix 210 and 215	Artic Technologies	$395-895	IBM	
Synphonix 240 and 245	Artic Technologies	$495-995	NEC	
Ufonic Voice System	Educational Technology	$245	Apple	
Vic-Talker/64-Talker	Talktronics, Inc.	$69.00	Commodore	
Votalker C-64	Votrox, Inc.	$59.95	Commodore	
Western Center Echo Syn Package	Western Center for Microcomputers in Spec. Ed.	$269	Apple, IBM	
Prose 4000	Speech Plus, Inc.	$1,750	IBM	
Accent-1200 and Accent-1600	Aicom Corp.	$625	Toshiba	
Accent-Mini	Aicom Corp.	$395	IBM	
Accent-PC	Aicom Corp.	$745	IBM	
Accent-SA	Aicom Corp.	$940-1440	IBM	
Syntha-Voice Models	Syntha Voice Computers, Inc.	$895	All	
Realvoice PC	Adaptive Communications Systems, Inc.	$1,595	IBM	
Sounding Board	GW Micro	$395	IBM, Toshiba	
Verbette Mark I	Computer Conversations, Inc.	$249.95	IBM	
Verbette Mark II	Computer Conversations, Inc.	$399.95	Multiple	

BEST COPY AVAILABLE

Table 3. Voice Output Computers

Brand Name	Manufacturer	Price	System	Description
Televox	Hexamedia	$1,895	IBM	Screen Review Program
Smoothtalker	First Byte,Inc.	$39.95-49.95	Multiple	Screen Review Program
Canon Print to Voice Computer 8020	Canon USA, Inc.	$4,250	All	
3278 Vert	TSI/VTEK	$495	All	
Voice Interactive Computer System	HyTek Manufacturing	$8,195-10,750	All	Voice Input Terminal
Notex	Adhoc Reading Systems, Inc.	$5,800	IBM	Braille Translator
Braille N Speak	Blazie Engineering	$905	IBM	Braille Translator
Talker II	Intex Micro Systems Corp.	$2,495	All	Computer Direct Selection Communicator
DragonDictate	Dragon Systems, Inc.	$9,000	IBM	Nonportable
Liaison	Du It Control Systems Group	$3,600-3,750	Apple, IBM	Nonportable
Nomad	Syntha Voice Computers, Inc.	$2,295	N/A	Portable
D'Light	Artic Technologies	$1,695-1,795	N/A	Portable
Eureka A4	Robotron Access Products, Inc.	$2,595	N/A	Portable
Keynote	Humanware, Inc.	$1,450-4,825	Apple/IBM	Portable
Laptalker and Laptalker Plus	Automated Functions, Inc.	$1,595-2,395	N/A	Portable

Table 4. Audio Output for Data Transmission

Brand Name	Manufacturer	Price	System	Description
Tweedle-Dump	John Monarch	$16.00	All	
Auditory Breakout Box	Smith Kettlewell Eye Research Inst.	$295.00	All	
WATCHDOG	Kansys, Inc.	$10.00	IBM	

Table 5. Large Print Hardware Systems

Brand Name	Manufacturer	Price	System	Description
Vantage CCD	TSI/VTEK	$2,795	IBM	Low Vision Computer Terminal Viewing System
Vu-Tek CRT Magnifier	Optical Devices	$199.95	Multiple	Monitor Screen Magnifier
Maxi-Screen	Engineering Consulting	$19.95	Apple	Monitor Screen Magnifier
Macnifier	Premtech Corp.	N/A	Apple	Monitor Screen Magnifier
NuVu Models MCI, PSI, and GNK	Less Gauss, Inc.	$144.95-179.95	All	Monitor Screen Magnifier
Anti-Glare Magnification Screen	Sher Mark Products, Inc.	$89.95	Apple	Monitor Screen Magnifier
Beamscope II	Florida New Concepts Marketing, Inc.	$68.95-78.95	All	Monitor Screen Magnifier
Compu-Lenz	Florida New Concepts Marketing, Inc.	$204.95	All	Monitor Screen Magnifier
Maclarger Video Monitor	Power R	$349	Apple	Large Print Display
Viking II, 2400 GS + 10 Monitors	Moniterm Corp.	$1,095-3,595	Apple, IBM	Large Print Display
Vista System, SFIB	TSI/VTEK	$2,095-2,495	IBM	Large Print Display
Visualtek DP-10	TSI/VTEK	$2,695	Apple	Large Print Display
Visualtek DP-11	TSI/VTEK	$1,995	IBM	Large Print Display
OPTEQ V and VFF	OPTEQ Vision Systems	$2,245-2,595	IBM	Large Print Display
Large Print Display Processor	TSI/VTEK	$2,295-2,895	Apple, IBM	Large Print Display
Vista and Vista 2	TSI/VTEK	$2,495	IBM	Large Print Display

Table 6. Large Print Software Programs

Brand Name	Manufacturer	Price	System	Description
Eye Relief	Skisoft Publishing Corp.	$295	IBM	Word Processing
Peachy Writer Bold	Cross Educational Software	$24.95	Apple	Text Editor
Magic Slate II	Sunburst Communications	$65-129	Apple	Word Processor
Multiscribe 3.0 and Multiscribe GS	Access Unlimited-Speech Enterprises	$79.95-99.95	Apple	Word Processor
B-Edit	Hexagon Products	$30.00	IBM	Word Processor
Large Character Tool Kit	Kidsview Software, Inc.	$49.95	Apple	Lesson Authoring
Large Print Word Processor	Benjamin Bayman	$20.00	Commodore	Word Processor
Qwerty Word Processor	HFK Software	$199-299	IBM	Voice Output Word Processor
Qwerty Forms Processor	HFK Software	$199-299	IBM	Voice Output Form Processor
Qwerty Large Print Reader	HFK Software	$99.00	IBM	Utility
Qwerty Large Print	HFK Software	$399	IBM	Graphics Utility
Tall Talk Prints 1	Access Unlimited-Speech Enterprises	$60	Apple	Printer Utility
Magic Slate Typestyles	Sunburst Communications	$49	Apple	Utility
Tall Talk Screens 1	Access Unlimited-Speech Enterprises	$45	Apple	Utility
B-Pop	Hexagon Products	$27	IBM	Utility
Big	Hexagon Products	$39	IBM	Utility for Lotus 1-2-3
In Focus	AI Squared	$149	IBM	Utility
Kidsword	Kidsview Software, Inc.	$39.95-49.95	Apple, Commodore	Word Processor
Kidsview	Kidsview Software, Inc.	$39.95	Commodore	Utility
Low Vision Editor (LVE) and LVE 23	Donald W. Ady	$20	TRS	Word Processor

Table 6. Large Print Software Programs (Continued)

Brand Name	Manufacturer	Price	System	Description
Magic Keyboard	Woodsmith Software Corp.	$44.50	IBM	Translation
PC Lens	Arts Computer Products, Inc.	$495	IBM	Utility
Talking Writer	Cross Educational Software	$24.95	Apple	Voice Output
BEX	Raised Dot Computing	$400	Apple	Voice Output Word Processor; Braille WP
Inlarge	Berkeley Systems, Inc.	$95	Apple	Voice Output
Large Print Word Processors	National Institute for Rehabilitation Engineering	$39.95-299.95	IBM	Word Processor
Spy Graf	LS & S Group	$295	IBM	Utility
Tall Talk Prints V.2	Access Unlimited-Speech Enterprises	$85	Apple	Voice Output
Tall Talk Screens V.2	Access Unlimited-Speech Enterprises	$65	Apple	Voice Output
Verbal View	Computer Conversations, Inc.	$249.95	Multiple	Utility
Zoomer	Kinetic Designs, Inc.	$130	IBM	Utility
Zoomtext	AI Squared	$495	IBM	Utility
LPDOS	Optelec US, Inc.	$495	IBM	Utility
Closeview	Apple Computer, Inc.	N/A	Apple	Utility
Handiview	Microsystems Software, Inc.	$195	Multiple	Utility
LPDOS Deluxe Edition	Optelec US, Inc.	$650	IBM	Utility

5.0 STATEMENT OF THE PROBLEM

As computers become more visually complex, new strategies are needed to augment the standard approaches to providing access to the information being displayed to persons with vision impairments in order to provide media access. Problems associated with computer input devices deal with the trouble of finding or identifying keys and controls on the keyboard and the problem of mouse driven control. Visually impaired individuals have difficulty using perfectly flat membrane keyboards, since they cannot find the keys even if they have memorized their position and function. They also have difficulty in locating keys on large keyboards without tactile landmarks. Visually impaired individuals cannot use a mouse because they cannot monitor the mouse cursor's continually changing position as they move.

Problems with computer output devices deal with the screen display and voice output. Some people with visual impairments cannot see lettering and symbols on keyboard, equipment or screen because it is too small or low contrast. They also need electronic access to information displayed on the screen in order to use special non-vision display substitutes. In other words, visually impaired individuals can use a portable voice output access device in place of the computer's standard screen display, except where these devices cannot get access to the contents (information) displayed on the computer's screen. Problems with computer input and output devices will become more severe in the future as computer systems move toward a more graphical approach to entering and displaying information.

The graphical user interface represents a fundamental change in the way computers display information and the way that humans interact with them. The most technical and fundamental difference is screen-rendering architecture. The use of pixels to put images on the screen leads to the problem of deciphering information on the screen. The second major difference involves the way that people interact with computers: how the graphical user interface represents information on the computer display and how users manipulate and control the flow of information. Some examples of current graphical applications which cause problems for the visually impaired include: icons, graphic charts, diagrams, pull-down menus, pop-up dialogue boxes, tear-off menus, pictures, animation, three dimensional images, and mouse-driven control.

6.0 THE DEPARTMENT OF EDUCATION'S PRESENT COMMITMENT AND INVESTMENT

According to the 1988 National Health Interview Survey, 600,000 Americans between the ages of 18 and 69 have blindness or visual impairments severe enough to limit their employment opportunities, and that number rises sharply with age. This is an indication of the size of the population who could potentially benefit from enhanced computer and electronic book access. Although the number of visually impaired people under 18 is relatively small, they can adapt to new computer access technologies most easily and use it for the rest of their lives.

With the advent of personal computers in 1975, the Department of Education began to fund research and development of computer input and output devices for sensory impaired people. Presently, the development of such devices is a stated research priority of the Department of Education as follows:

- The National Workshop on Rehabilitation Technology, sponsored by the Electronic Industries Foundation (EIF) and National Institute on Disabilities and Rehabilitation recommended making "information processing technology for access to print graphics, including computer access" top technology priority for visual impairments.

- Several of the funding criteria of the Department of Education's National Institute on Disability and Rehabilitation Research (NIDRR) are directed at the high unemployment rate of persons with vision impairments and severely visually impaired populations. Most severely visually impaired Americans are unemployed. Enhanced devices for computer access would improve the educational outlook of blind individuals, promote computer literacy, and improve employment opportunities and job retention among the computer literate. Another stated priority, advanced training for the blind and visually impaired at the pre- and post-doctoral levels, and in research, would benefit greatly from improved computer access technology.

- The Panel of Experts for the Department of Education program sponsoring this study consists of experts from industry and Government, including members of the sensory-impaired community. Their consensus opinion was that developing a larger Braille display is the highest priority for persons with visual impairments. Input and output devices for computer access ranked second. This rating was based on the lack of Braille devices and not the relative importance of the technologies or applications for all visually impaired individuals. However, the problem of computer input and output and electronic book access was considered crucial for media access and employment opportunities.

- One of the Department of Education's 1991 Small Business Innovative Research (SBIR) Program Research Topics is to develop or adapt communication devices for young children who are blind or deaf-blind.

The primary reason that electronic media access is a priority is that over two million persons with vision impairments could benefit from electronic information media access.

7.0 ACCESS TO PRINTED MEDIA INFORMATION MEDIA

Many federal, state, and local laws which influence access for persons with visual impairments. The most important single law related to access for persons who are vision impaired is Public Law 101-336, enacted July 26, 1990. Better known as the Americans with Disabilities Act (ADA), this law has broad implications for all disabled Americans and establishes the objective of providing access to persons with disabilities to physical and electronic facilities and media.

The other law that impacts technology for persons with visual impairments is Public Law 100-407-AUG 19, 1988 titled "Technology-Related Assistance for Individuals with Disabilities Act of 1988." Also known as the Tech Act, this law established a comprehensive program to provide for technology access to persons with disabilities. The law defines assistive technology devices: "Assistive technology devices means any item, piece of equipment, or product system, whether acquired commercially off the shelf, modified, or customized, that is used to increase, maintain, or improve functional capabilities of individuals with disabilities."

Computer access technology clearly meets this definition for persons with vision impairments and should be exploited to increase the ability of persons with vision impairments to obtain access to printed media. Within the findings and purpose of this law, computer access technology can provide persons with vision impairments with opportunities to:

- Exert greater control over their own lives by making computer literacy possible;
- Participate in and contribute more fully to activities in their home, school, and work environments, and in their communities; and
- Otherwise benefit from opportunities that are taken for granted by individuals who do not have disabilities.

8.0 POTENTIAL ACCESS IMPROVEMENTS WITH ADVANCED INPUT/OUTPUT DEVICES FOR COMPUTER & ELECTRONIC BOOK ACCESS TECHNOLOGY

This advanced media access technology offers the potential for dramatic improvements in information access for persons with vision impairments directly from their existing and future computer based information systems as follows:

- Databases
- Electronic mail systems
- Bulletin board systems
- Mail order systems
- Books and articles.
- Screen graphics

9.0 ADVANCED ELECTRONIC MEDIA TECHNOLOGIES

Several new technologies are emerging which will greatly improve graphical computer interface for the visually impaired. Computer input device technologies will attempt to solve the problems of mouse control and screen navigation in the absence of visual feedback and hand/eye coordination. Three such emerging technologies are the "UnMouse", handwriting-recognition systems, and CCD cameras. Voice recognition systems technology is also being pursued as a computer input device for the visually impaired. Computer output device technologies will attempt to provide alternate non-vision display substitutes through voice synthesizers and touch screens or enhanced images through head-up displays (HUDs). Table 7 depicts some of the advanced technology products currently available on the market.

Table 7. Advanced Technology Products

Brand Name	Manufacturer	Description
Pen Point	Go Corp.	Pen-input
Grid Pad	Grid Systems Corp.	Pen-input
Handwriter CAN	NCR Corp.	Pen-input
Telepad	Telepad Corp.	Pen-input
Glass Digitizer	Graphics Technology Co.	Glass digitizer with pen-input
Unmouse	Microtouch Systems, Inc.	Graphic input device
Video VGA	Truevision	Video
DS-3000	Chinon America, Inc.	Color Scanner
Wired for Sound	Aristosoft	Software to add sound to windows
Intouch	Berkeley Systems, Inc.	Software for tactile imaging device

The UnMouse is an input device designed to replace a mouse or trackball. It provides a means to relocate the cursor to a specific point on the screen and thus provides a reference location for a user to begin navigation. The touch-sensitive tablet combines three-input devices, providing cursor control, keypad functions, and stylus based graphics capabilities. The 3 inch by 4.5 inch tablet remains stationary beside the keyboard. Selection and cursor movement is done by sliding one's finger over the glass tablet and clicking by pressing on the tablet. Templates which slide under the glass or affix to the edge of the device labeled in braille, can be programmed and used as a function keypad. The UnMouse has an RS-232-C serial interface and is compatible with IBM PC, XT, AT, PS/2, and compatibles.

Handwriting recognition technology is an input device technology and could enhance computer access for the visually impaired. Through its application to pen-based computer systems, this technology will allow a visually impaired person to better interact with the computer by providing an alternate way to input information. The enabling technology for this emerging market is the incorporation of neural-network techniques into a flexible object-oriented operating system. System designers face several challenges including: creating a system that can adapt to multiple writers' handwriting, limiting the duration of system training, building a system that can recognize a wide enough range of characters, and allowing users to write naturally. The new technology will employ the following techniques to solve these challenges: examination of visual information; the handwritten text itself, analysis of data from the writing process, such as the sequence, timing, pressure and direction of pen strokes, and use of contextual data, such as predictable sequences of characters.

CCD Cameras could be utilized as computer input devices. They would work like a scanner but be more portable. The CCD Camera would use optical character recognition software to read screens, books, LCD, etc.

Voice recognition systems technology will allow visually impaired individuals to interface with the computer by way of voice input. In addition to enhancing normal interface with the computer, this technology would enable visually impaired people to accomplish data entry tasks. This technology encompasses everything from simple user-trained computer software and electronic circuits capable of recognizing a few single utterances to user adaptable continuous speech speaker-independent systems capable of 1,000 to over 20,000 words. Although the speaker-dependent systems have been on the market for over 10 years, the advanced technology speaker-adaptable continuous voice recognition systems are just beginning to make their appearance, and the speaker-independent continuous voice recognition systems are in research and development. These systems are expected to be available within 3 to 5 years for specific applications such as medical transcription.

The advanced technology voice recognition systems are using new computer digital signal processor boards, statistical software and advanced acoustic microphone technologies to achieve speaker adaptable, speaker independent continuous speech recognition systems that can recognize words, and form them into sentences in real time. One system under development from Dragon Systems Inc. uses an IBM PC 386 or 486 with a digital signal processor board and advanced statistical software to recognize over 10,000 words of single user adaptable natural speech. A list of organizations developing advanced technology voice recognition systems as well as further detailed information on this technology can be found in an alternate scenario entitled, "Voice Recognition Systems for Personal and Media Access."

Recently, programs of research in speech and natural language have been increasing in number and size all around the world. So far, there does not seem to be much divergence in the underlying technologies. However, there is an increasing divergence in goals. For instance, the European efforts see spoken dialogue systems as involving natural language generation and speech synthesis, as well as speech recognition and natural language understanding, while the DARPA community has generally seen the problem as "speech in, something else out;" thus there is little American effort on generation, and less on speech synthesis. An important difference in focus is that all the European efforts are multilingual in essence and by necessity, while most American work is on English only.

Voice Synthesizer technology has seen rapid growth recently especially in terms of improving the quality of the voice outputs. The focus is toward tailoring speech synthesis to the individual. By utilizing a smaller data base of words unique to a person, memory space and processing time can be reduced thus allowing for the possibility of a higher quality of voice output.

inTOUCH is a software utility program that allows TSI's Optacon II, a tactile imaging device, to display any information on the Apple Macintosh screen. Utilizing the mouse, this software allows visually impaired users to feel the Macintosh screen through a tactile array of vibrating pins. inTOUCH shows any part of the screen under the mouse. The user can scan lines of text at an adjustable rate of speed or move the mouse manually. Everything on the screen is accessible, including graphics.

Head-up displays could offer enhanced screen output images for the visually impaired. A HUD consists of three main parts. One component is a projector that emits the display light. Another element is a combiner that reflects the display light for creating the displayed images and also allows foreground objects to be seen. The third component is an electronic circuit that controls the display device and its brightness. HUDs were first installed in military airplanes in the 1940's, and are now widely used in commercial aircraft. They are also being considered for advanced automobile designs.

Aristosoft has developed a software utility program called "Wired for Sound" that lets users personalize their Windows desktop with over 50 voices, sound effects, and musical cues in a proprietary file format. Sounds can be assigned to application and icon messages, dialog boxes, and to the system startup and exit. Users can associate text in any dialog box with a specific sound. A talking clock lets you set alarms, and "Wired for Sound's" dynamic link library is included for use with macros in Windows applications, such as Excel. The cost is $49.

10.0 COST CONSIDERATIONS OF ADVANCED INPUT/OUTPUT DEVICE TECHNOLOGY

Table 8 shows the prices of the current advanced input/output device technology. Most of the technology is relatively new and thus prices have remained high. As competition increases, the cost of input/output device technology is expected to decrease as with other computer-related equipment. For example, as the second and third generation voice recognition products begin to appear, the cost of the technology will be

Table 8.

Penpoint	$4000
Grid Pad	$2500
Glass Digitizer	$ 400
Unmouse	$ 250
DS-3000	$1000
Personal Reader	$12,000

driven down by market forces and microelectronics implementations of voice recognition hardware.

Adapting certain advanced technologies for the purpose of enhanced computer access for the visually impaired may require a substantial investment that may not be practical for manufacturers to invest in without government assistance or sponsorship for the initial research and development phases. The reason for this is that the handicapped market is small which makes it more difficult to recover development costs within a production run without passing the full cost on to the consumer. The first applications are therefore usually systems adapted from mass market devices. With a systematic development approach to developing interfaces for applications for persons with visual impairments, the Department of Education can help reduce the cost of advanced input/output device technology to meet the needs of persons with visual impairments. This is possible because much of the research and development cost do not have to be amortized over the initial production runs.

11.0 COST BENEFITS TO PERSONS WITH SENSORY IMPAIRMENTS WITH EARLY INCLUSION OF SPECIAL ACCESS MODES

The cost benefits associated with early Department of Education sponsored research and development for application to persons with visual impairments is that the costs associated with this development will not have to be passed on to the user in the final product. The research and development areas for this targeted research should include: Vocabulary database development and structuring for voice recognition applications, interface requirement definition, human factors determination, and marketing and dissemination of information on potential uses. This will simplify integrating the needs of persons with visual impairments into the special access modes and reduce the development cost to manufacturers.

12.0 PRESENT GOVERNMENT INVOLVEMENT IN ADVANCED INPUT/OUTPUT TECHNOLOGY

The heaviest Government involvement in advanced input/output technology has been in the area of voice recognition. The U.S. Government involvement in voice recognition systems has been broad and includes National Security, Transportation, Commerce and Educational Applications. To date the most significant advanced technology effort is being conducted by DARPA's Information Science and Technology Office.

The U.S. Department of Education has maintained a large research program through both Grant and SBIR Programs for the past 30 years. Table 9 provides examples of programs currently active. The Department of Education programs provide the research and development platform essential to meet the computer input/output needs of persons with visual impairments. Without these programs to initiate new devices and probe new technologies, persons with sensory impairments would be denied access to advanced computer technologies.

Table 9. NIDRR Projects

Project	Organization
Rehabilitation Engineering Center on Access to Computers and Electronic Equipment	University of Wisconsin - Trace Center
A keyboard/voice interface for use of an on-line library system by the visually handicapped	Vatell Corporation
Screen manager for the IBM PC (SAM)	Automated Functions, Inc.
The Smith-Kettlewell Rehabilitation Engineering Center	Smith-Kettlewell Eye Research Foundation-Rehabilitation Engineering Center
Computer Access-Technology-Knowledge Base Expert System: Development, Evaluation, and Dissemination	Mississippi State University-RRTC on Blindness and Low Vision
A personal computer controller for multi-handicapped blind individuals	WesTest Engineering Corporation

The U.S. National Aeronautics and Space Administration (NASA) fund research and development efforts for devices for the handicapped through small business and university innovative research grant programs. Table 10 lists some of the current efforts which could aid persons with visual impairments.

Table 10. NASA Projects

Project	Organization
Optical Processing Technology	SBIR
Virtual Reality Head Set	Ames Research Center
Solid-State Laser Scanner	APA Optics, Inc.

13.0 ADVANCED INPUT/OUTPUT DEVICES FOR COMPUTER AND ELECTRONIC BOOK ACCESS TECHNOLOGY TIMELINE

Most of the advanced technologies for enhanced computer and electronic book access have had or will soon have first generation products on the market. For example, within the next one to two years, several user-independent continuous voice recognition systems are expected to be marketed based on the research sponsored by DARPA and private companies, such as, the America Telephone and Telegraph Corporation. Most of the advanced input/output technologies are expected to mature over the next five years to the point where they will provide computer control for persons with visual impairments. What are needed are comprehensive programs to apply the technologies to meet specific needs of persons with visual impairments. This will require that training programs be formulated and specific goals set to allow the input/output technologies to be adapted for use by persons with visual impairments.

A review of the Grant and SBIR programs should be conducted over the next two years to determine the most promising input/output devices to allow computer access. This review should provide a comprehensive list of priorities for future Grant and SBIR funding efforts. Following this review, the Department of Education should establish a program to fund 2 to 3 devices into the advanced development phase. This will allow a small business to implement the input/output device and help move it from the development phase to the production phase. This will ensure continued computer access for persons with vision impairments. Overall, the Grant and SBIR programs should be continued as structured to encourage the development of input/output devices for the visually impaired.

14.0 PROPOSED ROAD MAP FOR INCLUSION OF ELECTRONIC INFORMATION ACCESS CAPABILITIES

The Department of Education should begin the process of developing advanced input/output technology devices for computer and electronic book access for use by persons with visual impairments by developing several key technologies. Small Business Innovative Research (SBIR) Grants should be initiated in the areas of voice recognition, handwriting recognition, CCD cameras, speech synthesis, heads-up-displays and Braille technology.

These SBIR programs would consist of three phases. Phase I would involve concept studies and feasibility model development and would last approximately 6 months as presently structured. After a 6 month delay to resolve any outstanding issues, Phase II design would then last for approximately 18-24 months. A Phase III stage would be added to the SBIR process. Phase III would consist of Manufacturing Design and Analysis on input/output devices that offered the highest payoff to persons with visual impairments. The Department of Education would allow an Engineering Development Model to be built and fund approximately 20% of the initial grant to do the manufacturing analysis of the device. This phase will help alleviate the problem of the transition between research and production for small businesses. This would also involve providing assistance to small businesses in the form of recommendations and market size so they may be better qualified in attaining loans from the Small Business Administration.

The most promising programs from the SBIR's should be recommended for Innovative Grant programs. Field Initiated Grant programs should continue to be pursued when deemed appropriate. Because most of the technologies involved in this scenario are being developed for other commercial applications, 3-4 years seems a reasonable time period for each program. The payoff at the end of 3-4 years is the empowerment of persons with visual impairments to allow them to use systems that allow them equal access to computers and electronic books as well as access to personal communications services.

15.0 POTENTIAL PROGRAM SCHEDULE

Figure 1 is a proposed schedule for starting programs in advanced input/output device technology to meet the needs of persons with visual impairments.

	1992	1993	1994
Voice Recognition	X		
Handwriting Recognition		X	
CCD Cameras	X		
Speech Synthesis			X
Heads-Up Displays	X		
Braille Technology	X		

Figure 1.

In particular, the Department of Education needs to continue to identify specific needs and applications for advanced input/output device technology systems to meet the needs of persons with visual impairments. A comprehensive program would include the following:

- Description of the target audience
- Identification of specific needs
- Input techniques for computer applications
- Output techniques for computer applications
- Graphical user interfaces.

VISIBLE LIGHT SPECTRUM MANIPULATION TO ALLOW MEDIA ACCESS FOR PERSONS WITH SELECTIVE VISION

MARCH 1992

Prepared by

Daniel E. Hinton, Sr., Principal Investigator
and
Paolo Basso-Luca

SCIENCE APPLICATIONS INTERNATIONAL CORPORATION
3701 N. Fairfax Drive, Suite 1001
Arlington, VA 22203
(703) 351-7755

1.0 SCENARIO

Visible Light Spectrum Manipulation to Allow Media Access for Persons with Selective Vision.

2.0 CATEGORY OF IMPAIRMENTS

Persons with vision impairments.

3.0 TARGET AUDIENCE

Consumers with Vision Impairments. Persons with vision impairments will benefit from enhanced access to media information services and computer systems. This scenario on visible light spectrum manipulation technology provides a means to disseminate information to consumers with vision impairments. In particular, it provides a better understanding of the visible light spectrum manipulation technology available in electronic media over the next three to five years and the potential problems that could arise in media access.

Policy makers, including national representatives, government department heads, and special interest organizations. Policy makers will also benefit from this scenario because they can apply this scenario to better understand the issues related to electronic media access for persons with vision impairments. In addition, this scenario provides a point of departure for them to understand how advanced technology funding priorities with Government programs can accelerate access for persons with vision impairments to the ever expanding field of electronic media storage and retrieval of information. It will also provide a point of departure for legislation or regulatory action necessary to ensure electronic books and other electronic medias are accessible to persons with vision impairments.

Research and Developers. This group will benefit through a better understanding of the needs of persons with vision impairments and specifically their printed media communications requirements. This understanding of media access requirements will assist researchers and developers in designing media access functions in their future products to meet the needs of persons with impairments.

Manufacturers. Manufacturers will benefit through a better understanding of the potential market size and the existing Federal Government requirements for media access for persons with vision impairments which can be met by adding an access capability to their electronic media products.

4.0 THE TECHNOLOGY

Table 1 depicts some of the current products providing alternate displays which are available to persons with selective vision impairments to assist them in gaining media access.

Table 1. Alternate Display Systems Usable With All Software

Product Name	Vendor	Computer	Cost	Selective Vision Application
Advantage	Telesensory Systems Inc.	IBM	$31.95	Possible to have a positive or negative image on wither half of the split screen.
Anti-Glare Magnification Screen	Sher-Mark Products, Inc.	Apple	$89.95	Polarizing filter is used to reduce glare and improve contrast.
CloseView	Apple Computer, Inc.	Apple	N/A	Screen display can be changed from black on white to white on black.
inLARGE	Berkeley System Design	Apple	$95.00	Both black-on-white and white-on-black displays are possible.
Large Print Display Processor	VTEK	Apple, IBM	N/A	Screen image can be positive or negative.
PC Lens	Arts Computer Products, Inc.	IBM	$690.00	Color options
Spy Graf	LS & S Group	IBM	$295.00	Color options
Zoomer	Kinetic Designs, Inc.	IBM	$89.00	Color options

Advantage is a 19 inch version of Telesensory Systems Vantage closed circuit television system. Advantage provides large print access to a computer when used as a monitor for Vista. With Advantage, it is possible to have a positive or negative image on either half of the split screen, to block out all but one line of text, and to run two or more external monitors. Advantage also comes with a typewriter accessory that permits viewing of paper in a typewriter, and a horizontal and vertical underline/overline feature as an aid in reading text.

The Anti-Glare Magnification Screen fits over the screen of all Macintosh 128K, 512K, Plus, and SE computers. A magnification lens doubles the size of characters and images on the screen. In addition, a polarizing filter is used to reduce glare and improve contrast.

CloseView allows the screen on the Macintosh Plus, SE, and II to be magnified form two to sixteen times. Working as an option on the control panel, CloseView also permits the screen display to be changed from black on white to white on black. It is possible to turn CloseView on and off, and to change the magnification, by using combinations of the control, option, and other keys.

inLARGE is a software application which magnifies anything that normally appears on the Macintosh display by a factor of two to sixteen times. Characters as well as graphics are enlarged. inLARGE automatically follows the user's keystrokes and mouse movements without interfering with the application program. Visual cues let the user know where the cursor is located on the screen. Both black-on-white and white-on-black displays are possible.

Large Print Display Processor is a peripheral device which enlarges the text that appears on a computer's monitor screen. Text can be enlarged up to sixteen times its original size. The area to be enlarged can be moved around using a joystick control. The area enlarged can be made to follow the screen cursor. A single line may be isolated for viewing, and the screen image can be positive or negative. The Large Print Display Processor does not enlarge graphics.

PCLens is a program designed to make characters on a computer screen easier to read. Characters appear enlarged, spread apart, and colored (optional). The section of the normal display that appears enlarged on the screen may be moved about automatically or may be moved manually, both horizontally and vertically, using the cursor keys. PC Lens shows all the 255 IBM characters sent to the display screen.

Spy Graf provides memory-resident screen enlargement for IBM computers and compatibles that have at least 128K of memory. Characters on the screen can be enlarged up to 64 times normal size. If the program is run using a RGB card and monitor, any of sixteen screen colors may be selected. If used with a monochrome monitor, a monochrome graphics board is required.

Zoomer is a resident monitor enlargement program for IBM PC/XT/AT computers and true compatibles. The program can enlarge displays in text and graphics modes, as well as displays generated by CADICAM programs. The enlargement ranges from one to seven times. Movement of the Zoomer window can be accomplished by use of the keyboard or alternative input device. A scanning mode zoom window has the capabilities of user speed control and creation of cursor placement. Reverse video and color selection options are also included.

5.0 STATEMENT OF THE PROBLEM

Visual impairment is the second major cause of disability in the United States. Therefore, nearly every American who lives a full life eventually will be numbered among those who are considered "visually handicapped." Even men or women of forty require substantially more light when reading than their school-aged son or daughter. For many there is a constant deterioration of eyesight past middle-age, as shown in Figure 1. Approximately fifty percent of us, when we have passed the age of sixty, will have a detectable degree of cataract development, and once we are into our sixties, most of us will have significant difficulty in determining small details. Virtually all who reach the age of eighty or beyond will experience major deterioration in vision due to either inroads of maturity or disease.

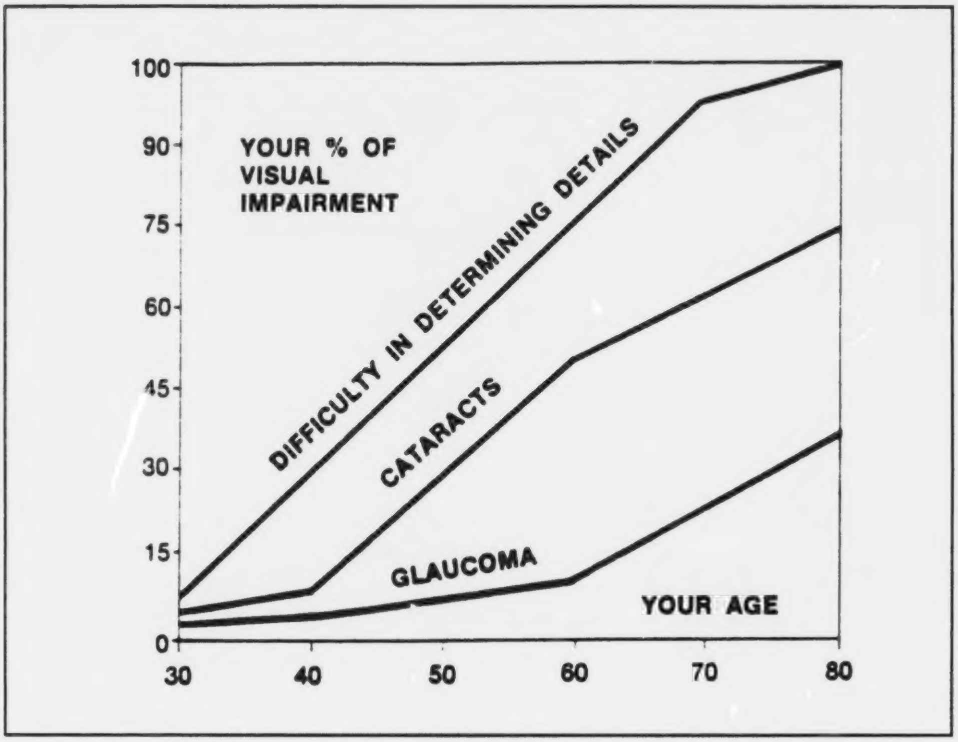


Figure 1.

Approximately 2.5 million Americans, many older than 65, suffer from low vision. Birth defects, injuries, and aging can cause low vision, but most cases are due to eye conditions that affect the retina, including:

- Macular Degeneration: Deterioration of the macula, the center of the retina used for sharp focus, causes central vision loss and makes reading difficult. This disability affects 20 percent of those over 75.

- Diabetic Retinopathy: Swelling and leakage of fluid in the center of the retina brought on by diabetes can cause scar tissue to form, leading to loss of sight.

- Glaucoma: Increased fluid pressure inside the eye damages the optic nerve, resulting in vision loss. Because peripheral or side vision usually is affected first, this disease sometimes is called "tunnel vision."

- Retinitis Pigmentosa: In this hereditary condition, the retina progressively deteriorates, causing loss of peripheral vision.

Cataracts (the clouding of the lens), cornea infections, and detachment of the retina also can cause low vision. In addition to loss of central or side vision, low vision patients may lose their color eyesight, have difficulty adapting to bright and dim light, and suffer diminished focusing power.

6.0 THE DEPARTMENT OF EDUCATION'S PRESENT COMMITMENT AND INVESTMENT

According to the 1988 National Health Interview Survey, 600,000 Americans between the ages of 18 and 69 have blindness or visual impairments severe enough to limit their employment opportunities, and that number rises sharply with age. This is an indication of the size of the population who could potentially benefit from enhanced computer and electronic book access. Although the number of visually impaired people under 18 is relatively small, they can easily adapt to new computer access technologies and use it for the rest of their lives.

With the advent of personal computers in 1975, the Department of Education began to fund research and development of computer input and output devices for sensory impaired people. Presently, the development of such devices is a stated research priority of the Department of Education as follows:

- The National Workshop on Rehabilitation Technology, sponsored by the Electronic Industries Foundation (EIF) and the National Institute on Disabilities and Rehabilitation recommended making "information processing technology for access to print graphics, including computer access" the top technology priority for visual impairments.

- Several of the funding criteria of the Department of Education's National Institute on Disability and Rehabilitation Research (NIDRR) are directed at the high unemployment rate of persons with vision impairments and severely visually impaired populations. Most severely visually impaired Americans are unemployed. Enhanced devices for computer access would improve the educational outlook of blind individuals, promote computer literacy, and improve employment opportunities and job retention among the computer literate. Another stated priority, advanced training for the blind and visually impaired at the pre- and post-doctoral levels, and in research, would benefit greatly from improved computer access technology.

- The Panel of Experts for the Department of Education program sponsoring this study consists of experts from industry and Government, including members of the sensory-impaired community. Their consensus opinion was that developing a larger Braille display is the highest priority for persons with visual impairments. Input and output devices for computer access ranked second. This rating was based on the lack of Braille devices and not the relative importance of the technologies or applications for all visually impaired individuals. However, the problem of computer input and output and electronic book access was considered crucial for media access and employment opportunities.

- One of the Department of Education's 1991 Small Business Innovative Research (SBIR) Program Research Topics is to develop or adapt communication devices for young children who are blind or deaf-blind.

The primary reason that electronic media access is a priority is that over two million people with vision impairments could benefit from electronic information media access.

7.0 ACCESS TO PRINTED MEDIA INFORMATION MEDIA

Many federal, state, and local laws influence access for persons with visual impairments. The most important single law related to access for persons who are vision impaired is Public Law 101-336, enacted July 26, 1990. Better known as the Americans with Disabilities Act (ADA), this law has broad implications for all disabled Americans and establishes the objective of providing access to persons with disabilities to physical and electronic facilities and media.

The other law that impacts technology for persons with visual impairments is Public Law 100-407-AUG 19, 1988 titled "Technology-Related Assistance for Individuals with Disabilities Act of 1988." Also known as the Tech Act, this law established a comprehensive program to provide for technology access to persons with disabilities. The law defines assistive technology devices: "Assistive technology devices means any item, piece of equipment, or product system, whether acquired commercially off the shelf, modified, or customized, that is used to increase, maintain, or improve functional capabilities of individuals with disabilities."

Computer access technology clearly meets this definition for persons with vision impairments and should be exploited to increase the ability of persons with vision impairments to obtain access to printed media. Within the findings and purpose of this law, computer access technology can provide persons with vision impairments with opportunities to:

- Exert greater control over their own lives by making computer literacy possible;

- Participate in and contribute more fully to activities in their home, school, and work environments, and in their communities; and

- Otherwise benefit from opportunities that are taken for granted by individuals who do not have disabilities.

8.0 POTENTIAL ACCESS IMPROVEMENTS WITH ADVANCED INPUT/OUTPUT DEVICES FOR COMPUTER AND ELECTRONIC BOOK ACCESS TECHNOLOGY

This advanced media access technology offers the potential for dramatic improvements in information access for persons with vision impairments directly from their existing and future computer based information systems as follows:

- Computer access

    -- Databases
    -- Electronic mail systems
    -- Bulletin board systems
    -- Mail order systems

- Books and articles.

9.0 ADVANCED ELECTRONIC MEDIA TECHNOLOGIES

9.1 Night Vision

Advanced night-vision equipment now under development by the U.S. military may soon find a home in a host of commercial applications, thanks to the Army's decision to declassify its uncooled thermal imaging sensor technology. The new devices could be utilized by the visually impaired as well as the normal population in automobiles as "vision enhancers" for nighttime drivers. Japanese, German and US automakers have toyed with the idea of automotive night vision devices for some time, but have been deterred by extremely high costs.

The new technology analyzes the temperature differences between an object and its background using infrared images that see through fog, smoke and dust to provide the viewer with a television-like picture. However, the new uncooled sensors are not burdened by cryogenic coolers, mechanical scanners, high-vacuum Dewars and high-pressure as bottles that current systems employ to cool the detector array. Instead, a new unscanned array with an integrated detector and readout structure that can be stabilized at room temperature will be used. Its temperature will be maintained by a simple one-stage thermoelectric cooler. For persons with retinitis pigmentosa or other disorders, smaller and lighter devices will be possible for use at night or in the daytime for setting the contrast of objects.

9.2 NASA Technology

Using space-based imaging techniques, NASA will develop a device designed to improve the eyesight of some 2.5 million Americans who suffer from low vision, a condition that cannot be corrected medically, surgically, or with prescription eyeglasses. The invention, called the Low Vision Enhancement System, will employ digital image processing technology. Experimenters will apply such processing techniques as magnification, spatial distortion, and contrast adjustment to compensate for blind spots in the patient's visual field.

The device will resemble mirrored wraparound sunglasses. Low-vision patients will view the outside world on color flat-panel television screens located in the lens portion of the glasses. Lenses and glass fibers will be embedded on each side of the wraparound section, where the front and ear pieces join. The lenses will transport images along the fibers to a miniature solid-state camera carried in a belt or shoulder pack. Images will be processed by a battery-powered computer in the pack and then transported via fiber back to the display screens for viewing.

The Low Vision Enhancement System should benefit patients with central vision loss, the part of vision normally used for reading. These patients may have macular degeneration associated with aging, or diabetic retinopathy, in which diabetes causes swelling and leakage of fluid in the center of the retina. It also could help patients with impaired side vision due to eye diseases such as retinitis pigmentosa.

10.0 COST CONSIDERATIONS OF ADVANCED TECHNOLOGY

Table 2 shows the prices of the current advanced light manipulation device technology. Some of the technology is relatively new and thus those prices have remained high. As competition increases, the cost of light manipulation device technology is expected to decrease as with other computer-related equipment.

Table 2.

Advantage	$31.95
Anti-glare Magnification Screen	$89.95
inLARGE	$95.00
PC Lens	$690.00
Spy Graf	$295.00
Zoomer	489.00

Adapting certain advanced technologies for the purpose of enhanced computer access for the visually impaired may require a substantial investment that may not be practical for manufacturers to invest in without government assistance or sponsorship for the initial research and development phases. The reason for this is that the handicapped market is small which makes it more difficult to recover development costs within a production run without passing the full cost on to the consumer. The first applications are therefore usually systems adapted from mass market devices. With a systematic development approach to developing interfaces for applications for persons with visual impairments, the Department of Education can help reduce the cost of advanced light manipulation device technology to meet the needs of persons with visual impairments. This is possible because much of the research and development cost do not have to be amortized over the initial production runs.

11.0 COST BENEFITS TO PERSONS WITH SENSORY IMPAIRMENTS WITH EARLY INCLUSION OF SPECIAL ACCESS MODES

The cost benefits associated with early Department of Education sponsored research and development for application to persons with visual impairments is that the costs associated with this development will not have to be passed on to the user in the final product. The research and development areas for this targeted research should include: interface requirement definition, human factors determination, and marketing and dissemination of information on potential uses. This will simplify integrating the needs of persons with visual impairments into the special access modes and reduce the development cost to manufacturers.

12.0 PRESENT GOVERNMENT INVOLVEMENT IN ADVANCED TECHNOLOGY

The U.S. Department of Education has maintained a large research program through both Grant and SBIR Programs for the past 30 years. Table 3 provides examples of programs currently active. The Department of Education programs provide the research nd development platform essential to meet the computer input/output needs of persons with visual impairments. Without these programs to initiate new devices and probe new technologies, persons with sensory impairments would be denied access to advanced computer technologies.

The U.S. National Aeronautics and Space Administration (NASA) fund research and development efforts for devices for the handicapped through small business and university innovative research grant programs. Table 4 lists some of the current efforts which could aid persons with visual impairments.

Table 3. NIDRR Projects

Project	Organization
Rehabilitation Engineering Center on Access to Computers and Electronic Equipment	University of Wisconsin - Trace Center
New Techniques for Low Vision	Smith-Kettlewell Eye Research Institute
The Smith-Kettlewell Rehabilitation Engineering Center	Smith-Kettlewell Eye Research Foundation-Rehabilitation Engineering Center
Computer Access-Technology-Knowledge Base Expert System: Development, Evaluation, and Dissemination	Mississippi State University-RRTC on Blindness and Low Vision

Table 4. NASA Projects

Project	Organization
Virtual Reality Head Set	Ames Research Center
Contrast Adjustment for Maculopathy	Jet Propulsion Laboratory

13.0 ADVANCED TECHNOLOGY TIMELINE

Most of the advanced technologies for enhanced computer and electronic book access have had or will soon have first generation products on the market. Most of the advanced light manipulation technologies are expected to mature over the next five years to the point where they will provide computer access for persons with visual impairments. What are needed are comprehensive programs to apply the technologies to meet specific needs of persons with visual impairments. This will require that training programs be formulated and specific goals set to allow the light manipulation technologies to be adapted for use by persons with visual impairments.

A review of the Grant and SBIR programs should be conducted over the next two years to determine the most promising light manipulation devices to allow computer access. This review should provide a comprehensive list of priorities for future Grant and SBIR funding efforts. Following this review, the Department of Education should establish a program to fund two to three devices into the advanced development phase. This will allow a small business to implement the light manipulator device and help move it from the development phase to the production phase. This will ensure continued computer access for persons with vision impairments. Overall, the Grant and SBIR programs should be continued as structured to encourage the development of light manipulation devices for the visually impaired.

14.0 PROPOSED ROAD MAP FOR INCLUSION OF ELECTRONIC INFORMATION ACCESS CAPABILITIES

The Department of Education should begin the process of developing advanced light manipulation technology devices for computer and electronic book access for use by persons with visual impairments by developing several key technologies. Small Business Innovative Research (SBIR) Grants should be initiated in the areas of Infrared Sensors, Digital Image Processing, CCD cameras, and heads-up-displays.

These SBIR programs would consist of three phases. Phase I would involve concept studies and feasibility model development and would last approximately 6 months as presently structured. After a 6 month delay to resolve any outstanding issues, Phase II design would then last for approximately 18-24 months. A Phase III stage would be added to the SBIR process. Phase III would consist of Manufacturing Design and Analysis on input/output devices that offered the highest payoff to persons with visual impairments. The Department of Education would allow an Engineering Development Model to be built and fund approximately 20% of the initial grant to do the manufacturing analysis of the device. This phase will help alleviate the problem of the transition between research and production for small businesses. This would also involve providing assistance to small businesses in the form of recommendations and market size so they may be better qualified in attaining loans from the Small Business Administration.

The most promising programs from the SBIR's should be recommended for Innovative Grant programs. Field Initiated Grant programs should continue to be pursued when deemed appropriate. Because most of the technologies involved in this scenario are being developed for other commercial applications, 3-4 years seems a reasonable time period for each program. The payoff at the end of 3-4 years is the empowerment of persons with visual impairments to allow them to use systems that allow them equal access to computers and electronic books as well as access to personal communications services.

15.0 POTENTIAL PROGRAM SCHEDULE

Figure 3 is a proposed schedule for starting programs in advanced input/output device technology to meet the needs of persons with visual impairments.

In particular, the Department of Education needs to continue to identify specific needs and applications for advanced light manipulation device technology systems to meet the needs of persons with visual impairments. A comprehensive program would include the following:

	1992	1993	1994
Infrared Sensors		X	
CCD Cameras	X		
Digital Image Processing			X
Heads-Up Displays	X		

Figure 3.

- Description of the target audience
- Identification of specific needs
- Output techniques for computer applications
- Graphical user interfaces issues.

FLAT PANEL TERMINAL DISPLAYS USED WITH PAGE SCANNERS

MARCH 1992

Prepared by

Daniel E. Hinton, Sr., Principal Investigator
and
Paolo Basso-Luca

SCIENCE APPLICATIONS INTERNATIONAL CORPORATION
3701 N. Fairfax Drive, Suite 1001
Arlington, VA 22203

1.0 SCENARIO

Flat Panel Terminal Displays Used with Page Scanners

- Character Readers for Dynamic LED and LCD Display Access

2.0 CATEGORY OF IMPAIRMENTS

Persons with vision impairments.

3.0 TARGET AUDIENCE

Consumers with Vision Impairments. Persons with vision impairments will benefit from enhanced access to media information services and computer systems. This scenario on flat panel terminal displays used with page scanners provides a means to disseminate information to consumers with vision impairments. In particular, it provides a better understanding of the optical character recognition system technology available in electronic media over the next three to five years and the potential problems that could arise in media access.

Policy makers, including national representatives, Government department heads, and special interest organizations. Policy makers will also benefit from this scenario because they can apply this scenario to better understand the issues related to electronic media access for persons with vision impairments. In addition, this scenario provides a point of departure for them to understand how advanced technology funding priorities within Government programs can accelerate access for persons with vision impairments to the ever expanding field of electronic media storage and retrieval of information. It will also provide a point of departure for legislation or regulatory action necessary to ensure electronic books and other electronic medias are accessible to persons with vision impairments.

Researchers and Developers. This group will benefit through a better understanding of the needs of persons with vision impairments and specifically their printed media communications requirements. This understanding of media access requirements will assist researchers and developers in designing media access functions in their future products to meet the needs of persons with vision impairments.

Manufacturers. Manufacturers will benefit through a better understanding of the potential market size and the existing Federal Government requirements for media access for persons with vision impairments which can be met by adding an access capability to their electronic media products.

4.0 THE TECHNOLOGY

4.1 Flat-CRT Technology

Some observers think that the new flat-CRT technology is a viable way of producing CRT displays that could compete with LCDs for future computer laptop and other flat-panel applications. Such CRTs would be as think and lightweight as LCDs, but brighter, less power hungry and cheaper to make. A vacuum-microelectronics display uses thousands of minute cone-shaped cathodes called microtips. They emit a stream of electrons that jump across a small vacuum gap toward a phosphor-coated anode to create images.

4.2 Optical Character Recognition

Early versions of OCR software used a matrix-matching technology to recognize characters which involved storing pattern templates for every type of style and size that might appear in a scanned document. The program would then attempt to match each character it scanned against the resident images. In addition to consuming enormous amounts of memory, the software required the power of a coprocessor mounted on an add-in board, which cost upward of $2000, on machines with a 286 or lower CPU.

Adequate processing power without hardware assistance became available with the introduction of 386-class CPUs, but the increase in laser-printer documents presented another technological hurdle to OCR. The matrix-matching technique sufficed for monofont, typewritten documents, but it could not cope with the profusion of typefaces and sizes found in many laser-printed documents and faxes. A typical document contains far more characters because laser printers are not limited to the fixed pitch of typewriters, and you can now kern text and use proportionally spaced typefaces.

Consequently, current products now incorporate some form of automatic character recognition based on topological feature extraction, which consists of an algorithm that extracts salient features of each character and compares them to each other. Some of the programs still use a form of matrix technology to aid in the recognition process. Table 1 provides a representative list of optical scanners on the market today.

The basic technology of a flatbed scanner is relatively simple. Within a sealed box, a fluorescent or incandescent light bulb illuminates the image to be scanned (called the target), and a photosensor called a CCD (charged coupled device) absorbs the target's reflected light. The CCD is essentially an array of thousands of light-detecting cells, each of which produces a voltage level in proportion to the amount of light it picks up. An analog-to-digital converter then processes these voltages into digital values, whose precision is based on the number of bits per pixel supported by the scanner. On an 8-bit scanner, the range of brightness levels that the CCD can "see" on the target can be divided into 256 shades of gray. Because of the limitations of CCD technology, most page scanners do not really capture a full 8 bits of usable information; electronic noise reduces the actual

Table 1. Optical Character Recognition Devices

Brand Name	Manufacturer	Price	System	Description
PC/KPR	Kurzweil Imaging Systems, Inc.	$3,995-6,995	IBM	OCR system with voice output
Personal Reader (KPR)	Kurzweil Imaging Systems, Inc.	$7,950-11,950	All	OCR system with voice output
Kurzweil 5000, 5100 and 5200 Scanning Systems	Kurzweil Imaging Systems, Inc.	N/A	IBM	OCR system with voice output
Adhoc Reader	Adhoc Reading Systems, Inc.	$6,290	IBM	OCR system with voice output
Arkenstone Reader Models S and E	Arkenstone	$1,495-3,995	IBM	OCR system with voice output
Cannon 1X 12 Scanner, Cannon PC Interface Board and Readright V1.13 Software	Canon USA, Inc.	Scanner $795 Software $595 Board $395	IBM	OCR system
Oscar	TSI/VTEK	$3,895-4,295	IBM	OCR system
Discover 7320 Models 10, 20, 30	Kurzweil Imaging Systems, Inc.	$3,995-6,995	IBM	OCR system
Omni-Reader	IMPX	$199	Apple, IBM	OCR scanner
Totec Model TO-5050 Proscan and TO-5000B	Totec Co. Ltd. Legal Scan Serve, Inc.	$9,990	N/A	OCR scanner
PC Scan 1020 and 2020	Dest Corp.	$1,900-1,945	Apple, IBM	OCR scanner
Deskscan 2000 and 3000	Chinon America, Inc.	N/A	Apple, IBM	OCR scanner
Personal Computer Scanner (PCS)	Compuscan, Inc.	$3,495	IBM	OCR scanner
Scan 300/S	Abaton Technology Corp.	$1,595	Apple, IBM	OCR scanner
Readstar II Plus	Inovatic	$995	IBM, Apple	OCR software
Readright 2.0	OCR Systems	$495	IBM	OCR software
Docread I, III and Expert	Adhoc Reading Systems, Inc.	$2,690-6,290	IBM	OCR software for Adhoc Reader
Read-It	Olduvai Corp.	$295-595	Apple	OCR software

resolvability of the image to 7 or even 6 bits. Once the scanner creates the image, a high-speed direct-interface card transmits the image to the PC.

To capture the color information, they make three passes, successively shining light through red, green, and blue filters. Eight bits of information are recorded for each color channel to give you 24-bit color.

Because of their limitations, these devices are not a suitable replacement for full-page desktop scanners. Most hand-held scanners can scan only a bit more than 4 inches in width in a single pass, although large images can be pieced together with multiple scans. Also, because most hand-held scanners are manually dragged across the image being scanned, image quality depends on how the user moves the scanner. The smoother and straighter the movement, the better the quality of the resulting scanned image.

Synthesized speech is one of the most powerful and least expensive access devices for the blind. Generally, a speech system consists of resident software that converts text into speech, a speech-synthesis board with audio amplification and an interface to the PC bus, and a speaker that sits outside the computer. When users optically read a series of text, the system turns the letters into phonemes (the smallest units of sound), runs through a series of rules that tell it how to say the word, and outputs the word through the external speaker. Tables 2-4 show the numerous speech and audio products on the market designed specifically for persons with vision impairments.

5.0 STATEMENT OF THE PROBLEM

As computers become more visually complex, new strategies are needed to augment the standard approaches to providing access to the information being displayed to persons with vision impairments in order to provide media access. The problem associated with computer input deals with character readers for dynamic LED and LCD flat panel terminal display access. The problem with computer output deals with voice output. This scenario attempts to point out the key technologies which can be utilized in solving the following problem process:

- The user scans a text-based document from a flat panel terminal display into a PC using a hand-held or flatbed scanner.

- OCR software running on the PC "recognizes" bit-mapped characters in documents generated by the computer terminal.

- Some packages must be manually "trained" by the users to read new text. Other packages read any type automatically.

Table 2. Speech Synthesizers

Brand Name	Manufacturer	Price	System	Description
Doubletalk	RC Systems, Inc.	$249.95	Apple, IBM	
Apollo	Dolphin Systems	$687	IBM	
Readme System; Termivox; Termiscreen Reader	Infonox	$1695 $1995 $445	N/A	
Echo+ Speech Synth	Street Electronics Corp.	$119.95-179.95	Apple	
Votrax, Personal Speech System	Votrax	$449.95	Apple, IBM	Voice output module
Accent-MC	Aicom Corp.	N/A	IBM	
Accent-XE	Aicom Corp.	N/A	Toshiba	
Synphonix 230 and 235	Artic Technologies	$595-1,095	Toshiba	
Synphonix 310 and 315	Artic Techologies	$695-1,095	IBM	
Synphonix 250 and 255	Artic Technologies	$695-1,195	Toshiba	
Echo Commander	Street Electronics Corp.	$164.19	Apple	
Synphonix 220 and 225	Artic Technologies	$495-995	Toshiba	
DECTALK	Digital Equipment Corp.	$4,498	All	
Intex-Talker	Intex Micro Systems Corp.	$345	All	Voice output module
Echo II	Street Electronics Corp.	$116.95	Apple	
Artic Crystal	Artic Technologies	$1,195-2,095	IBM	
Audapter Speech System	Personal Data Systems, Inc.	$1,095	All	
Blackboard	Peripheral Technologies, Inc.	$595	Apple	
Calltext 5000	Centigram	$3,225	IBM	
Calltext 5050	Speech Plus, Inc.	$3,900	All	
Echo 1000	Street Electronics Corp.	$134.95	Tandy	
Echo IIC	Street Electronics Corp.	$134.95	Apple	

BEST COPY AVAILABLE

Table 2. Speech Synthesizers (Continued)

Brand Name	Manufacturer	Price	System	Description
Echo MC	Street Electronics Corp.	$179.95	IBM	
Echo PC+	Street Electronics Corp.	$161.95	IBM	
Personal Speech System	Votrax, Inc.	$449	All	
Syntha-Voice Model I	Syntha Voice Computers, Inc.	$695	IBM	
Speaqualizer	American Printing House for the Blind	$809.41	IBM	
Speech Adapter for PC Convertible	IBM Corp.	$620	IBM	
Speech Thing	Covox, Inc.	$79.95	IBM	
Synphonix 210 and 215	Artic Technologies	$395-895	IBM	
Synphonix 240 and 245	Artic Technologies	$495-995	NEC	
Ufonic Voice System	Educational Technology	$245	Apple	
Vic-Talker/64-Talker	Talktronics, Inc.	$69.00	Commodore	
Votalker C-64	Votrox, Inc.	$59.95	Commodore	
Western Center Echo Syn Package	Western Center for Microcomputers in Spec. Ed.	$269	Apple, IBM	
Prose 4000	Speech Plus, Inc.	$1,750	IBM	
Accent-1200 and Accent-1600	Aicom Corp.	$625	Toshiba	
Accent-Mini	Aicom Corp.	$395	IBM	
Accent-PC	Aicom Corp.	$745	IBM	
Accent-SA	Aicom Corp.	$940-1440	IBM	
Syntha-Voice Models	Syntha Voice Computers, Inc.	$895	All	
Realvoice PC	Adaptive Communications Systems, Inc.	$1,595	IBM	
Sounding Board	GW Micro	$395	IBM, Toshiba	
Verbette Mark I	Computer Conversations, Inc.	$249.95	IBM	
Verbette Mark II	Computer Conversations, Inc.	$399.95	Multiple	

BEST COPY AVAILABLE

Table 3. Voice Output Computers

Brand Name	Manufacturer	Price	System	Description
Televox	Hexamedia	$1,895	IBM	Screen Review Program
Smoothtalker	First Byte,Inc.	$39.95-49.95	Multiple	Screen Review Program
Canon Print to Voice Computer 8020	Canon USA, Inc.	$4,250	All	
3278 Vert	TSI/VTEK	$495	All	
Voice Interactive Computer System	HyTek Manufacturing	$8,195-10,750	All	Voice Input Terminal
Notex	Adhoc Reading Systems, Inc.	$5,800	IBM	Braille Translator
Braille N Speak	Blazie Engineering	$905	IBM	Braille Translator
Talker II	Intex Micro Systems Corp.	$2,495	All	Computer Direct Selection Communicator
DragonDictate	Dragon Systems, Inc.	$9,000	IBM	Nonportable
Liaison	Du It Control Systems Group	$3,600-3,750	Apple, IBM	Nonportable
Nomad	Syntha Voice Computers, Inc.	$2,295	N/A	Portable
D'Light	Artic Technologies	$1,695-1,795	N/A	Portable
Eureka A4	Robotron Access Products, Inc.	$2,595	N/A	Portable
Keynote	Humanware, Inc.	$1,450-4,825	Apple/IBM	Portable
Laptalker and Laptalker Plus	Automated Functions, Inc.	$1,595-2,395	N/A	Portable

Table 4. Audio Output for Data Transmission

Brand Name	Manufacturer	Price	System	Description
Tweedle-Dump	John Monarch	$16.00	All	
Auditory Breakout Box	Smith Kettlewell Eye Research Inst.	$295.00	All	
WATCHDOG	Kansys, Inc.	$10.00	IBM	

- Once the OCR software recognizes the bit-mapped characters, it translates them into a variety of text file formats, including ASCII and formats used by specific word processing programs. The files can then be called up from within word-processing or desktop publishing applications.

- A speech synthesizer package provides voice output capability.

6.0 DEPARTMENT OF EDUCATION'S PRESENT COMMITMENT AND INVESTMENT

According to the 1988 National Health Interview Survey, 600,000 Americans between the ages of 18 and 69 have blindness or visual impairments severe enough to limit their employment opportunities, and that number rises sharply with age. This is an indication of the size of the population who could potentially benefit from enhanced computer access. Although the number of visually impaired people under 18 is relatively small, they can adapt to new computer access technologies most easily and use it for the rest of their lives.

With the advent of personal computers in 1975, the Department of Education began to fund research and development of computer input and output devices for sensory impaired people. Presently, the development of such devices is a stated research priority of the Department of Education as follows:

- The National Workshop on Rehabilitation Technology, sponsored by the Electronic Industries Foundation (EIF) and the National Institute on Disabilities and Rehabilitation recommended making "information processing technology for access to print graphics, including computer access" the top technology priority for visual impairments.

- Several of the funding criteria of the Department of Education's National Institute on Disability and Rehabilitation Research (NIDRR) are directed at the high unemployment rate of persons with vision impairments and severely visually impaired populations. Most severely visually impaired Americans are unemployed. Enhanced devices for computer access would improve the educational outlook of blind individuals, promote computer literacy, and improve employment opportunities and job retention among the computer literate. Another stated priority, advanced training for the blind and visually impaired at the pre- and post-doctoral levels, and in research, would benefit greatly from improved computer access technology.

- The Panel of Experts for the Department of Education program sponsoring this study consists of experts from industry and Government, including members of the sensory-impaired community. Their consensus opinion was that developing a larger Braille display is the highest priority for persons with visual impairments. Input and output devices for computer access ranked second. This rating was based on the lack of Braille devices and not the relative importance of the technologies or applications for all visually impaired individuals. However, the problem of computer input and output was considered crucial for media access and employment opportunities.

- One of the Department of Education's 1991 Small Business Innovative Research (SBIR) Program Research Topics is to develop or adapt communication devices for young children who are blind or deaf-blind.

The primary reason that electronic media access is a priority is that over two million persons with vision impairments could benefit from electronic information media access.

7.0 ACCESS TO PRINTED MEDIA INFORMATION MEDIA

Many federal, state, and local laws influence access for persons with visual impairments. The most important single law related to access for persons who are vision impaired is Public Law 101-336, enacted July 26, 1990. Better known as the Americans with Disabilities Act (ADA), this law has broad implications for all disabled Americans and establishes the objective of providing access to persons with disabilities to physical and electronic facilities and media.

The other laws that impacts technology for persons with visual impairments is Public Law 100-407-AUG 19, 1988 titled "Technology-Related Assistance for Individuals with Disabilities Act of 1988." Also known as the Tech Act, this law established a comprehensive program to provide for technology access to persons with disabilities. This law defines assistive technology devices: "Assistive technology devices means any item, piece of equipment, or product system, whether acquired commercially off the shelf, modified, or customized, that is used to increase, maintain, or improve functional capabilities of individuals with disabilities."

Computer access technology clearly meets this definition for persons with vision impairments and should be exploited to increase the ability of persons with vision impairments to obtain access to printed media. Within the findings and purpose of this laws, computer access technology can provide persons with vision impairments with opportunities to:

- Exert greater control over their own lives by making computer literacy possible;
- Participate in and contribute more fully to activities in their home, school, and work environments, and in their communities; and
- Otherwise benefit from opportunities that are taken for granted by individuals who do not have disabilities.

8.0 POTENTIAL ACCESS IMPROVEMENTS WITH ADVANCED INPUT/OUTPUT DEVICES FOR COMPUTER AND ELECTRONIC BOOK ACCESS TECHNOLOGY

This advanced media access technology offers the potential for dramatic improvements in information access for persons with vision impairments directly from their existing and future computer based information systems as follows:

- Databases
- Electronic mail systems
- Bulletin board systems
- Mail order systems
- Books and articles.

9.0 ADVANCED ELECTRONIC MEDIA TECHNOLOGIES

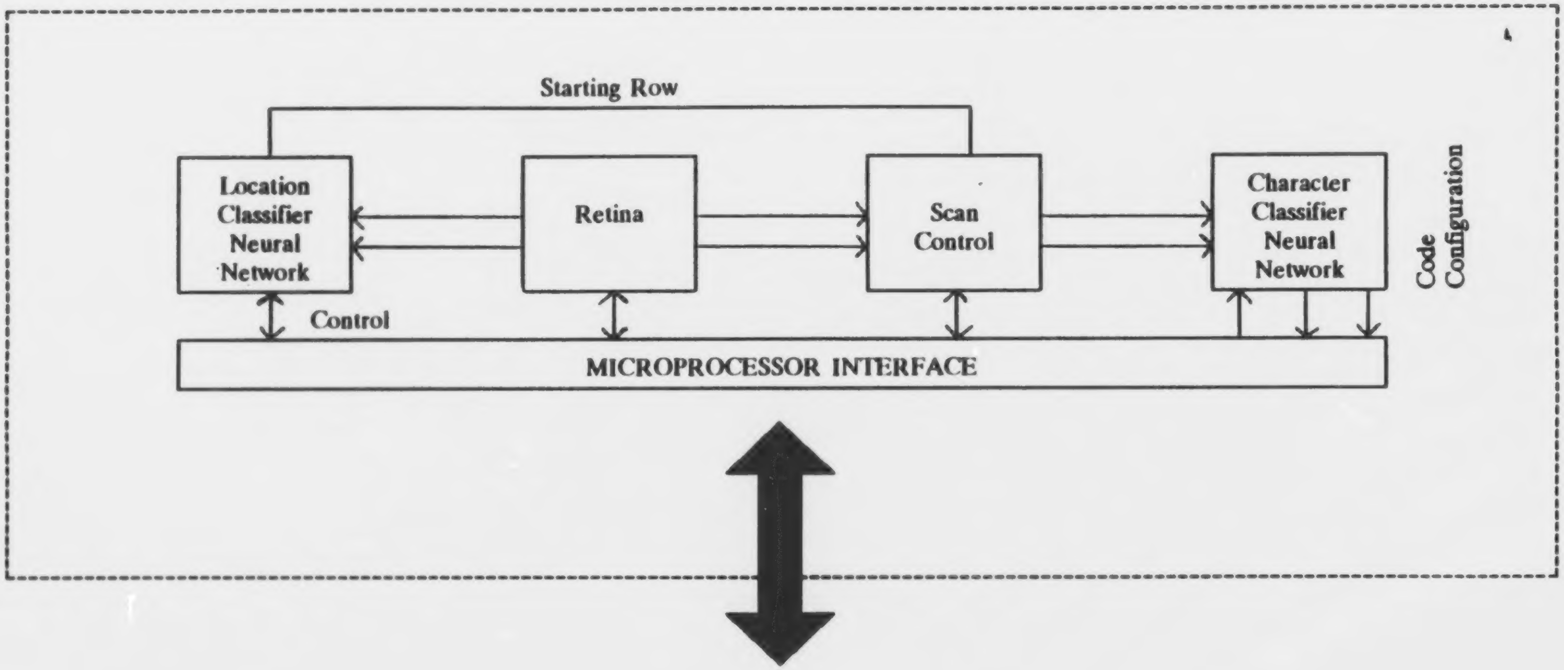
Several new technologies are emerging which will enhance access for the visually impaired to flat panel terminal displays using optical character recognition.

Synaptics, Inc. has developed an optical character recognition system that it says is faster than existing solutions because image sensing and classification are performed in parallel on a single piece of silicon. The OCR chip packs an analog sensing array, two neural networks and a digital controller and extracts analog functionality from its digital circuitry.

The new OCR chip operation is modeled on the human eye and ear that use digital circuitry to perform analog functions. A block diagram of the system is shown in Figure 1.

While conventional OCR systems require an expensive, high-bandwidth connection between the sensor and the computer's memory, the Synaptics approach eliminates the dependence on high-speed off-chip communication pathways. If you want to perform high-speed optical recognition, you are limited by the bandwidth between the sensor and the classifier-just 30 images per second, using a TV camera. But putting the sensor on the same chip with the classification circuitry allows the chip to do the same task thousands of times per second.

The OCR chip has a two-dimensional, 60x20-pixel sensing array. Sensing an image and segmenting it out into a recognizable character take only 1 microsecond. Assembly of an appropriate binary code shows throughput to 2000 characters/second, still an order of magnitude faster than conventional approaches. Key to the systems' speed is that sensing and classification are performed in parallel. All sensors pass their values onto the classifier simultaneously using hundreds of parallel connections. The chip operates much more closely to the biological model of the eye than anything else that is available. The sensing section is modeled on the retina, and the on-chip high-bandwidth connection to the classifier is modeled on the optic nerve's connection between the eye and the brain.



Starting Row

Location Classifier Neural Network

Retina

Scan Control

Character Classifier Neural Network

Code Configuration

Control

MICROPROCESSOR INTERFACE

- Parallel performance of sensing and classification allows high speed.

Figure 1. OCR System Modeled on Human Eye, Ear

UMAX Technologies has developed a standalone OCR machine called ReadStation which combines a scanner, automatic document feeder, dedicated computer, and Caere Corporation's OmniPage OCR software. Printed or typewritten documents are fed into the ReadStation, converted to electronic form, and written as files to the built-in 3.5 inch disc drive. Word processing, spreadsheet, and database file formats such as WordPerfect, Lotus 1-2-3, and dBase are supported, and selectable using a control panel on the front of the unit. The unit can be connected to a PC via an RS-232C or RS-422 serial port interface for direct file transfer.

In text mode, graphics and images are automatically ignored or removed, and settings can be adjusted to read specific areas of a page. In graphics mode, images are saved to a TIFF file format. ReadStation accommodates a maximum document size of 8.5x14 inches and has a maximum recognition rate of 115 characters per second.

CCD cameras could be utilized as computer input devices. They would work like a scanner but be more portable. The CCD camera would use optical character recognition software to read screens, books, or LCD to name a few examples.

Handwriting recognition technology could also be tied in with optical character recognition to enhance visually impaired access to handwritten materials. This technology will allow a visually impaired person to be able to read mail, handwritten notes, etc. with little or no assistance. The enabling technology for this emerging market is the incorporation of neural-network techniques into a flexible object-oriented operating system. Pen-input computers are of little or no direct benefit to most of the severely visually impaired population, but their development has reawakened interest in handwriting recognition recently. System designers face several challenges, including: creating a system that can adapt to multiple writers handwriting, limiting the duration of system training, building a system that can recognize a wide enough range of characters, and allowing users to write naturally. The new technology will employ the following techniques to solve these challenges: examination of visual information; the handwritten text itself, analysis of data from the writing process, such as the sequence, timing, pressure and direction of pen strokes, and use of contextual data, such as predictable sequences of characters. Scanned handwritten text contains no time and pressure information, but recognizing it is otherwise analogous to recognizing text on a pen-input computer. Companies which are developing handwriting recognition systems are listed in Table 5.

Voice synthesizer technology has seen rapid growth recently, especially in terms of improving the quality of the voice outputs. The focus is toward tailoring speech synthesis to the individual. By utilizing a smaller database of words unique to a person, memory space and processing time can be reduced, thus allowing for the possibility of a higher quality of voice output.

Table 5. Companies Developing Handwriting Recognition Systems

Grid Systems, Inc.	Freemont, CA
Go Corporation	Foster City, CA
Microsoft Corporation	Redmond, WA
Momenta Corporation	Santa Clara, CA
Nestor, Inc.	Providence, RI
Active Book Co., Ltd.	Cambridge, England

10.0 COST CONSIDERATIONS OF ADVANCED TECHNOLOGY

Table 6 shows the prices of the current advanced optical character recognition device technology. Most of the technology is relatively new and thus prices have remained high. As competition increases, the cost of OCR device technology is expected to decrease as with other computer-related equipment. For example, as the second and third generation products begin to appear, the cost of the technology will be driven down by market forces and microelectronics implementations of OCR hardware.

Table 6. Current Prices of OCR Devices

PC/KPR	$7000
Adhoc Reader	$6500
Arkenstone Reader, Models S and E	$4000
Oscar	$4500
DS-3000	$1000
Personal Reader	$12,000

Adapting certain advanced technologies for the purpose of enhanced computer access for the visually impaired may require a substantial investment that may not be practical for manufacturers to invest in without government assistance or sponsorship for the initial research and development phases. The reason for this is that the handicapped market is small which makes it more difficult to recover development costs within a production run without passing the full cost on to the consumer. The first applications are therefore usually systems adapted from mass market devices. With a systematic development approach to developing interfaces for applications for persons with visual impairments, the Department of Education can help reduce the cost of advanced OCR device technology to meet the needs of persons with visual impairments. This is possible because much of the research and development cost do not have to be amortized over the initial production runs.

11.0 COST BENEFITS TO PERSONS WITH SENSORY IMPAIRMENTS WITH EARLY INCLUSION OF SPECIAL ACCESS MODES

The cost benefits associated with early Department of Education sponsored research and development for application to persons with visual impairments is that the costs associated with this development will not have to be passed on to the user in the final product. The research and development areas for this targeted research should include: Vocabulary database development and structuring for voice synthesis, interface requirement definition, human factors determination, and marketing and dissemination of information on potential uses. This will simplify integrating the needs of persons with visual impairments into the special access modes and reduce development cost to manufacturers.

12.0 PRESENT GOVERNMENT INVOLVEMENT IN ADVANCED TECHNOLOGY

The U.S. Department of Education has maintained a large research program through both Grant and SBIR Programs for the past 30 years. Table 7 provides examples of programs currently active. The Department of Education programs provide the research and development platform essential to meet the computer input/output needs of persons with visual impairments. Without these programs to initiate new devices and probe new technologies, persons with sensory impairments would be denied access to advanced computer technologies.

Table 7. NIDRR Projects

Project	Organization
Rehabilitation Engineering Center on Access to Computers and Electronic Equipment	University of Wisconsin - Trace Center
The Smith-Kettlewell Rehabilitation Engineering Center	Smith-Kettlewell Eye Research Foundation-Rehabilitation Engineering Center
Computer Access-Technology-Knowledge Base Expert System: Development, Evaluation, and Dissemination	Mississippi State University-RRTC on Blindness and Low Vision
A personal computer controller for multi-handicapped blind individuals	WesTest Engineering Corporation

The U.S. National Aeronautics and Space Administration (NASA) fund research and development efforts for devices for the handicapped through small business and university innovative research grant programs. Table 8 lists some of the current efforts which could aid persons with visual impairments.

Table 8. NASA Projects

Project	Organization
Optical Processing Technology	SBIR
Solid-State Laser Scanner	APA Optics, Inc.

13.0 ADVANCED TECHNOLOGY TIMELINE

Most of the advanced technologies for optical character recognition have had or will soon have first generation products on the market. Most of the advanced input/output technologies are expected to mature over the next five years to the point where they will provide computer control for persons with visual impairments. What are needed are comprehensive programs to apply the technologies to meet specific needs of persons with visual impairments. This will require that training programs be formulated and specific goals set to allow the OCR technologies to be adapted for use by persons with visual impairments.

A review of the Grant and SBIR programs should be conducted over the next two years to determine the most promising OCR devices to allow computer access. This review should provide a comprehensive list of priorities for future Grant and SBIR funding efforts. Following this review, the Department of Education should establish a program to fund 2 to 3 devices into the advanced development phase. This will allow a small business to implement the OCR device and help move it from the development phase to the production phase. This will ensure continued computer access for persons with vision impairments. Overall, the Grant and SBIR programs should be continued as structured to encourage the development of OCR devices for the visually impaired.

14.0 PROPOSED ROAD MAP FOR INCLUSION OF ELECTRONIC INFORMATION ACCESS CAPABILITIES

The Department of Education should begin the process of developing advanced OCR technology devices for use by persons with visual impairments by developing several key technologies. Small Business Innovative Research (SBIR) Grants should be initiated in the areas of flat panel terminal displays, handwriting recognition, CCD cameras, and speech synthesis.

These SBIR programs would consist of three phases. Phase I would involve concept studies and feasibility model development and would last approximately 6 months as presently structured. After a 6 month delay to resolve any outstanding issues, Phase II design would then last for approximately 18-24 months. A Phase III stage would be added to the SBIR process. Phase III would consist of Manufacturing Design and Analysis on OCR devices that offered the highest payoff to persons with visual impairments. The Department of Education would allow an Engineering Development Model to be built and fund approximately 20% of the initial grant to do the manufacturing analysis of the device. This phase will help alleviate the problem of the transition between research and production for small businesses. This would also involve providing assistance to small businesses in the form of recommendations and market size so they may be better qualified in attaining loans from the Small Business Administration.

The most promising programs from the SBIR's should be recommended for Innovative Grant programs. Field Initiated Grant programs should continue to be pursued when deemed appropriate. Because most of the technologies involved in this scenario are being developed for other commercial applications, 3-4 years seems a reasonable time period for each program. The payoff at the end of 3-4 years is the empowerment of persons with visual impairments to allow them to use systems that allow them equal access to computers as well as access to printed media.

15.0 POTENTIAL PROGRAM SCHEDULE

Figure 2 is a proposed schedule for starting programs in advanced OCR device technology to meet the needs of persons with visual impairments.

	1992	1993	1994
Handwriting Recognition			X
CCD Cameras	X		
Speech Synthesis		X	
Flat Panel Terminal Displays		X	

Figure 2. Proposed Schedule

In particular, the Department of Education needs to continue to identify specific needs and applications for advanced OCR device technology systems to meet the needs of persons with visual impairments. A comprehensive program would include the following:

- Description of the target audience
- Identification of specific needs
- Input device development
- Software development
- Voice output.

DESCRIPTIVE VIDEO FOR TELEVISION ACCESS

MARCH 1992

Prepared by

Daniel E. Hinton, Sr., Principal Investigator
and
Charles Connolly and Lewis On

SCIENCE APPLICATIONS INTERNATIONAL CORPORATION
3701 N. Fairfax Drive, Suite 1001
Arlington, VA 22203

1.0 SCENARIO

Described Video for Television Access.

2.0 CATEGORY OF IMPAIRMENTS

Persons with vision impairments.

3.0 TARGET AUDIENCE

Consumers with Vision Impairments. Persons with vision impairments will benefit from enhanced access to television media services. This scenario on described video provides a means to disseminate information to consumers with vision impairments. In particular, it provides a better understanding of the technology available to provide described video to persons with vision impairments over the next three to five years.

Policy makers, including national representatives, Government department heads, and special interest organizations. Policy makers can use this scenario to better understand the issues related to television media access for persons with vision impairments. In addition, it provides a point of departure for policy makers to understand how advanced technology legislative, regulatory, and funding priorities within Government programs can accelerate development of described video for persons with vision impairments.

Researchers and Developers. This group will benefit through a better understanding of the needs of persons with vision impairments and specifically their television media access needs. Understanding media access requirements will assist researchers and developers in designing access to described video into their future products to meet the needs of persons with vision impairments.

Manufacturers and Broadcasters. Manufacturers and broadcasters will benefit through a better understanding of the requirements of described video, the potential market size and the existing Federal Government requirements for television media access for persons with vision impairments which can be met by adding described video capability to their systems.

4.0 THE TECHNOLOGY

This report is based on information from many sources, which are listed in the bibliography, but three sources deserve special acknowledgement: Station WGBH, Boston; Smith-Kettlewell Eye Research Institute; and COSMOS Corporation. WGBH, as the leading producer of described video, provided information on production and distribution of described video on PBS. Smith-Kettlewell provided a report, commissioned by the U.S. Department of Education, entitled "Technical Viability of Descriptive Video Services, June 1990," and COSMOS Corporation provided a report, also commissioned by the U.S. Department of Education, entitled "Commercial Viability of Descriptive Video Services, May 1990." For brevity, these will be referred to as the Smith-Kettlewell report and the COSMOS report, respectively.

This scenario does not necessarily reflect the views of the U.S. Government or the U.S. Department of Education, and the mention of specific products and trade names does not imply their endorsement by the U.S. Government or the U.S. Department of Education.

4.1 Introduction

A visually impaired person in front of a television has limited access to information that is only presented visually. Described video (DV) uses narration to describe the essential features of what is happening on the television screen, omitting anything that is clear from the sound track alone. Video description can be anything from spontaneous comments to the scripted narration produced by a few small private TV networks, up to the carefully-developed scripted narration produced by WGBH Boston's Descriptive Video Service Department. "Descriptive Video Service" and "DVS" are service marks of WGBH Boston, and no endorsement of WGBH or its DVS Department is intended by reference to them.

Described video was inspired by theater description. In 1981, the Washington Ear began narrating theater productions for the visually impaired, using infrared technology to transmit the narration to members of the audience with special receivers.

In 1985, the Public Broadcasting Service (PBS) adopted the Multichannel Television Sound (MTS) System, making it possible for PBS to experiment with an additional TV audio channel for narration. The Federal Communications Commission (FCC) had already made MTS the protected standard for multichannel television sound the year before, but networks had not taken advantage of it yet. The MTS system defines four simultaneously broadcast audio channels per television station.[1] Those channels are the stereo sum, stereo difference, second audio program (SAP), and professional (Pro) channels. In technical terms, the added channels are subcarriers of the stereo sum (main) audio channel. These added channels are ignored by non-stereo TVs and VCRs. All televisions, VCRs and TV radios use the stereo sum channel for TV sound. Stereo televisions have the added capability of combining the stereo sum and stereo difference channels, if both are present, to produce stereo sound. Many stations in major metropolitan areas broadcast in stereo; but not all stations are equipped to broadcast in stereo. The SAP channel is accessible with most stereo televisions and stereo VCRs, but few non-PBS stations are equipped to broadcast on the SAP channel. (The Pro channel, which will be discussed later in the scenario, was intended for station use, such as sending cues to remote TV crews, so only a few of the most expensive televisions can receive the Pro channel.)

---

[1] The MTS system, developed by Zenith, plus noise reduction developed by DBX, makes up the Broadcast Television Systems Committee (BTSC) system. In practice, the MTS system is always used with DBXs noise reduction system, so this scenario will refer to the BTSC system as the MTS system, emphasizing the multichannel aspect of it.

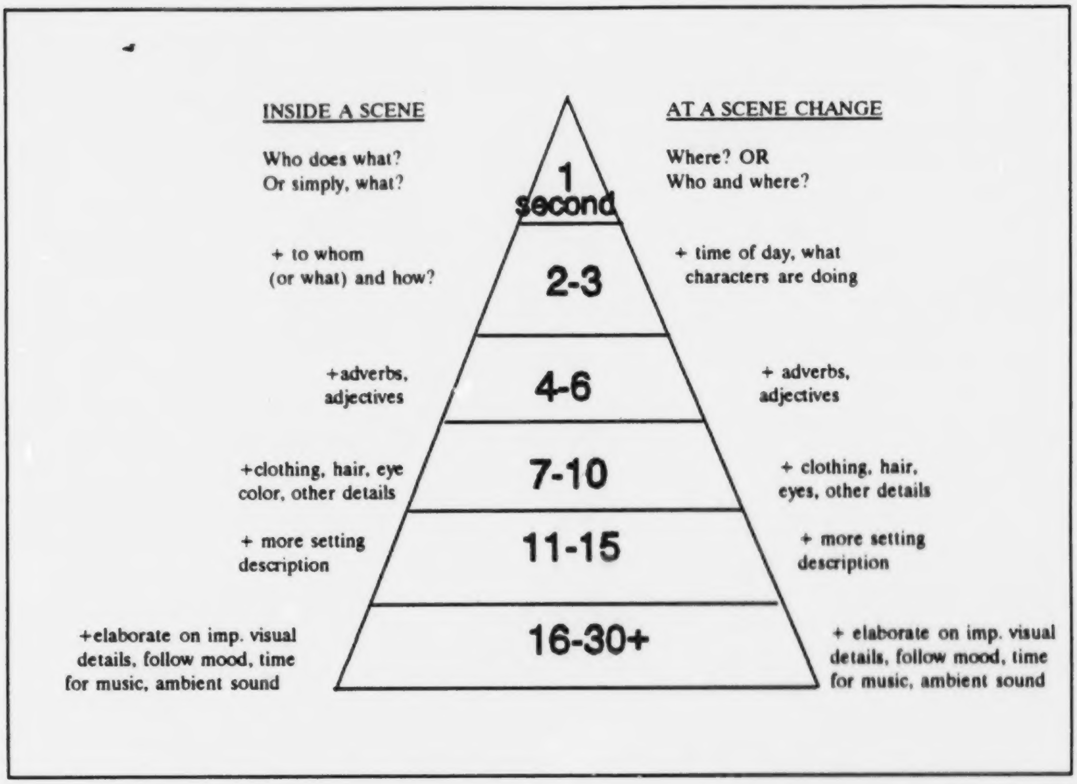
In 1986, PBS station WGBH, Boston, started experimenting with described video using the second audio program (SAP) channel. The SAP channel was chosen because it is widely accessible but independent of the stereo sum and difference channels; no one hears the narration unless they select the SAP channel. Anyone with a stereo TV or stereo VCR can access the SAP channel of the station they are watching by pushing a button or two on their remote control.

PBS began regularly scheduled described video broadcasts on the SAP channel in January 1990, with 28 stations participating. That number has more than doubled in two years, and PBS stations are still adding SAP capability as they modernize and upgrade.

4.2 DVS Production

An example of Described Video (DV) narration, from the WGBH DVS Style Manual, may be helpful: "The scene changes to an outdoor circus at dusk. Milo throws peanuts at a clown." None of that information would be clear from the standard audio track alone because there would be no accompanying dialogue; however, the setting is often essential to understanding events and dialogue. Throwing peanuts at a clown may suggest a carefree mood, an angry child, etc. Ideally, described video objectively sketches what is on the screen; interpretation is left to the listener. There is a limit as to how far that is possible, however, because a detailed description of everything on the screen would interrupt the dialog and bore the listener. Prioritizing is essential. The choice of what to point out is inevitably somewhat subjective, so producing high quality narration requires devoting a great deal of effort to creating narration that is as faithful to the original production as possible. Figure 1, from the WGBH *DVS Style Manual*, shows the prioritization used by WGBH.

As implemented by WGBH, the narration sound track is mixed with the original program sound track for broadcast on the SAP channel. That procedure is followed because most stereo televisions and stereo VCRs do not provide the option of listening to both the SAP channel and the standard audio program at the same time. That would be the more flexible option in the future, however, allowing each listener to adjust the relative volume of the sound track and narration. Unfortunately, listening to stereo plus SAP would require an extra audio expander in the receiver; the cost is very small but the TV market is so price-sensitive that it is not expected to be added unless customers demand it. If the main audio program could be omitted from the video description channel, it would also reduce the cost of producing described video, though not dramatically. Mixing the regular sound track with the narration takes an extra production step: adjusting the relative volume of the two audio tracks.



INSIDE A SCENE
AT A SCENE CHANGE
Who does what? Or simply, what?
1 second
Where? OR Who and where?
+ to whom (or what) and how?
2-3
+ time of day, what characters are doing
+adverbs, adjectives
4-6
+ adverbs, adjectives
+clothing, hair, eye color, other details
7-10
+ clothing, hair, eyes, other details
+ more setting description
11-15
+ more setting description
+elaborate on imp. visual details, follow mood, time for music, ambient sound
16-30+
+ elaborate on imp. visual details, follow mood, time for music, ambient sound

Figure 1. Prioritization in DVS Production

After the producer agrees in writing to have the program described, WGBH uses the following sequence to produce DV for a program: the producer sends a tape of the program to PBS six weeks before air time and a copy is sent to DVS at WGBH. The "describer" listens to the program, given little or no access to the picture, to find out where narration is needed. Then, a special program, running on a Macintosh SE/30 computer, helps to note where narration can be placed in the program without talking over dialog or important sounds in the original sound track. WGBH points out that less computer assistance is possible, but the equipment cost is more than made up by the reduced labor cost and elimination of tedious work. The computer program assumes a narration rate of 180 words per minute. Based on the information gathered with the program, a script is produced and edited twice. Two priorities must be balanced to make the script both informative and aesthetically pleasing: scenes and visual events should be described dynamically, thoroughly, and objectively, but excessive detail must be avoided. Also, information that is clear from the original sound track should not be repeated in the narration. Finally, the narration is read while watching and listening to the show, and the narration is mixed with the original sound track. Then, the narration can be dubbed onto the tape for broadcast.

WGBH has found that approximately 16-20 hours of "describer" time and 2.5 hours of narration-time are needed for each program hour. Equipment required includes the computer and custom program, a TV monitor, a 3/4-inch videocassette recorder, and a VCR interface, such as the ARTI box for Society of Motion Picture and Television Engineers (SMPTE) time code. A simple sound studio is also required.

5.0 STATEMENT OF THE PROBLEM

5.1 Demand for Described Video

Most Americans rely on television for entertainment. Television has also become an important source of information, and not just in the form of news and documentaries. The medium of television has had a profound effect on our culture: the way we see ourselves and the outside world.

DV enables a visually impaired person to share the experience of watching television with friends and family. Access does not depend on the patience, age and describing skills of friends and relatives. For example, few adults, let alone children, could adequately describe the events in Shakespeare's "Hamlet." Fewer still would be able to convey the costumes, gestures and settings. With described video, it makes no difference if the listener is with friends or home alone. An elderly person who is visually impaired may want to watch television with friends who are elderly and visually impaired; described video can make that easy, but it also lets visually impaired people be more independent when they want to be. All viewers get high quality narration without dividing anyone's attention. Since the narrative is planned in advance, the narrator never talks over the dialog, has to retract misinterpretations, or repeats anything that is self-evident from the dialog. Although DV is a new medium, it can be compared with the radio dramas of the 1920's, 30's and 40's. Instead of relying on a picture on a screen to convey visual information, the radio dramas painted pictures for the mind's eye.

Entertainment options for people with severe visual impairments are often limited. Many severe visual impairments make getting to places like movie theaters and playhouses difficult; for example, driving a car may be impossible. Fifty-five percent of the severely visually impaired population are age 75 or older, compounding the problem of access to outside entertainment. See Table 1, from the COSMOS report, for the age distribution of the severely visually impaired population. Many visually impaired people, especially those who are elderly, have a fixed income. Most blind people are unemployed, and many people with visual impairments are underemployed. The cost of tickets to movies, plays, concerts and sporting events makes them inaccessible to much of the visually impaired population, and other access issues, such as seating and lighting, compound the problem. Described television can provide a relatively inexpensive form of entertainment to these people; often the only entertainment available. The COSMOS study points out that severe visual impairments usually set in after age 44, so people are likely to be used to watching television by then and continue to do so.

Table 1. Estimated Population with Severe Visual Impairment by Age, 1986

Age in Years	Population (% of impaired population)			
	Estimate 1*		Estimate 2**	
Under 18	38,300	(2%)	38,300	(1%)
18-44	137,300	(7%)	137,300	(5%)
45-64	273,400	(14%)	273,400	(10%)
65-74	384,800	(20%)	822,000	(29%)
75 & Over	1,100,000	(57%)	1,552,700	(55%)
ALL AGES	1,933,800	(100%)	2,823,700	(100%)

* Estimate 1 uses 1977 national HIS rates, the latest year this measure was asked of the whole sample. The rates were applied to 1986 general population estimates by the U.S. Bureau of the Census.

**Estimate 2 uses 1984 rates which came from a special HIS supplement for old persons only (65+), with improved interview techniques.

Having a visual impairment is not required for one to benefit from DV. Virtually everyone can benefit from it, but the benefit is greatest to persons who are visually impaired and/or using a SAP-equipped radio instead of a TV. WGBH estimated that 11.5 million people with visual impairments can benefit from DV, which is the approximate size of the visually impaired population, as shown in Table 2, from the COSMOS study. After watching several high-quality described programs, it is easy to understand why that estimate is reasonable; DV keeps the viewer from missing important visual details that make programs more interesting and easier to follow. This feature may also make DV useful for people with attention disorders. A person who uses DV does not need to self identify and/or register as blind or visually impaired or buy any specialized equipment. There are no prescriptions, no forms to fill out, and no one else has to know about it. Privacy and easy access help to insure that people who may benefit from DV will try it. That is important, because a service must be used to be useful.

Not every program should have description with it, however. Turner Broadcasting's Cable News Network (CNN), for example, is made for television but broadcast over radio in many major cities, including Los Angeles and Washington, DC. That arrangement works for several reasons: news programs are highly written, there is little or no time for detailed description, and to begin with, the script is usually "written to the picture." This is not the type of program most requested by visually impaired viewers. Furthermore, parts of a

Table 2. Potential DVS Users by Level of Visual Impairment

Level of Visual Impairment	Estimated Population	Source	Date
Totally Blind--no or little sensitivity to light	0.05 million	American Foundation for the Blind (AFB)	1978
Legally Blind--acuity of 20/200 or worse in better eye with correction or a visual field of 20 degrees diameter or less	0.6 million	AFB	1986
Severely Visually Impaired--inability to read newsprint with corrective lenses	1.4 million	National Society to Prevent Blindness (NSPB)	1980
Severely Visually Impaired--inability to read newsprint with corrective lenses or, if under six years old, blind in both eyes, or having no useful vision in either eye	1.9-2.8 million	AFB	1986
Same as above; augmented by AFB's estimate of 500,000 institutionalized	2.4-3.3 million	AFB	1989
Visually Impaired--chronic or permanent defect resulting from disease, injury, or congenital malformation that results in trouble seeing in one or both eyes even when wearing glasses	8.4 million	National Center for Health Statistics (NCHS)	1988
Visually Impaired--same as above, includes color blindness, vision in only one eye, and other non-severe problems	12 million	WGBH testimony	1989

newscast are virtually non-stop talking, so there would be few gaps to be filled by narration, and the narration would sometimes have to be spontaneous for a live broadcast. Spontaneous narration is inherently of lower quality than pre-planned narration.

Generally the degree of dramatic or emotional content, plus the relative importance of visual information, indicate how useful it will be to describe a program. The relevance of the program to people with visual impairments may also be a consideration, but that can be difficult to estimate.

Commercial networks sometimes receive programs 6 or 8 hours before broadcast; PBS has a policy of six weeks lead time, though they have added video description in 36-44 hours at least once before. The COSMOS study points out that it can be hard to get producers to submit shows for captioning, which may also cause delays in producing described video.

5.2 Commercial Networks' Concerns About Distributing Described Video to Affiliates

Producing and distributing described video demands careful planning and special equipment. The COSMOS study found that even the major commercial networks feel that DV is "the right thing to do," but they are unable or unwilling to invest millions of dollars to produce and distribute an extra audio track for programs without some assurance that it will attract about a million new viewing households. Commercial networks are willing to add a new service if it improves their Neilson rating by at least one point, attracting 1% of the viewing audience.

The commercial networks' experience with closed captioning, which has never had the audience many people feel it merits, tends to reinforce their fears that DV could become another important but underutilized service. A closer look raises serious questions about the analogy. Visual impairment and hearing impairment influence people's lives very differently. Also, DV over the SAP channel is accessible with any off-the-shelf stereo TV or VCR; closed captioning has required a special decoder that is associated with hearing impairment. Special decoders reduce audience size, which is why closed captioning decoders will be built into all new TVs starting in 1993. Finally, as implemented, closed captioning appears as writing on a TV screen, which does not lend itself to portability. DV, on the other hand, is sound. It could easily be incorporated into Walkman-style radios and car radios, for example, potentially enjoying a wide audience and all of the cost and feature benefits that come with it. These issues are discussed further in Section 10.0 of this scenario.

There is also a fear that DV could cause problems with the automatic switching that has become standard for the major networks, because their switching equipment does not provide for an extra sound track. That means manually switching the extra sound track for programs that offer video description. All stations are concerned about switching and routing errors because they might cause a loss of audio broadcasting time on the order of seconds.

Distribution from Network to Affiliates

According to the Smith-Kettlewell report, distributing described video from a network to its affiliate stations via satellite is relatively straightforward and relatively inexpensive. Satellites should have enough extra bandwidth to handle an extra audio channel until the vertical blanking interval (VBI) technology, described later in this scenario, takes over in a few years. At the network end, an extra audio subcarrier would cost about $250 to $1050 in equipment. The lower cost would be to add a module to the network's "audio subcarrier processor;" the higher cost would be to add another uplink channel processor if the existing one had no vacant slots for modules. The network would only incur this cost once. At the affiliate end, the same cost range, $250 to $1050, would apply for modules for an "audio subcarrier processor," but each affiliate would need the equipment to receive DV and incur the one-time equipment cost.

Modifications to Affiliates

To transmit DV, network affiliates would have to route the narration track from the satellite downlink to the station's transmitter. The transmitter and antenna are generally located together and within 15 miles of the affiliate station. An audio line would be required from the "audio subcarrier processor" module to an optional simple audio console for the added audio channel. The next step would depend on how the station is configured. Stations normally use a microwave link to send their programs from the station to the transmitter, but some stations send the programs to the transmitter with subcarriers; others add the subcarriers at the transmitter. If the subcarriers are added at the station, the audio console would go straight to the SAP generator and that would be all. Otherwise, the station would have to add a module at each end of the microwave link to accommodate the extra audio channel. According to the Smith-Kettlewell report, the audio console would cost anywhere from $2,000 for a used console to $10,000 for a new one. The microwave link modules run about $2,100 a pair, and SAP generators are about $4,500 to $8,000. Some stations would need all of these items, others would already have some or all of them.

Transmitter Requirements

The Smith-Kettlewell report found that pre-MTS transmitters may not permit SAP transmission for various reasons, but they are also difficult to maintain. Therefore, most of them will have been replaced within five years. Some of the replacement transmitters come with built-in SAP and Pro generators. Replacing a transmitter would presumably be a good time to add SAP capability, since Smith-Kettlewell indicates "the main cost of rewiring the station facility would be in labor and would depend entirely on how the facility was constructed." The FCC could require all stations to add the SAP channel when upgrading their facilities. If that were to happen, all stations would be SAP-capable within 10 to 15 years.

Cable TV Issues

Smith-Kettlewell reports that the cost of bringing DV to cable TV depends upon how cable systems are set up. "Block conversion" systems handle transmission on the VBI, which will be discussed later in this scenario, and subcarriers, such as the added MTS channels, without modification. These systems are also the most common arrangement at the cable station end. "Base-band processing" systems, on the other hand, make scrambling and other features easier to implement, but they have problems with subcarriers and VBI. At the subscriber end, "conversion might include changing [the] subscribers' boxes," depending on the type of system used at the cable station end. No cost estimate was given for the conversion.

The real policy issue with cable companies is assuring that a program related service like DV (and captioning) is passed through to the cable subscriber and not stripped off at the cable headend. Federal cable legislation that includes language requiring this is pending.

5.3 Present Availability of Described Video

As of early 1992, DV is still experimental, but the U.S. Department of Education funds video description for a growing list of PBS series: "American Playhouse," "The American Experience," "Wild America," "Masterpiece Theater," and "Nature." The National Endowment for the Arts and matching funds raised by WGBH allow the described video production of "Mystery!," "Degrassi High," and "The WonderWorks Family Movie." The Nostalgia Television Network, which is not affiliated with PBS, broadcasts "classic" movies with video description on the main audio channel. Everyone must listen to narration on the main audio channel, because it is inseparable from the main sound track. If most listeners want the narration, there is no problem, but a major network does not have that option for its programs. Other sources of DV include Metropolitan Washington Ear, Audio Optics, Inc., and Audio Vision.

5.4 Technologies Used to Transmit DVS from Affiliate Stations Now

Table 3 summarizes the technologies covered in this section (present technologies) and in Section 9.0 (advanced technologies), for transmitting DV to homes. The present technologies are: the stereo sum (main audio) channel, the stereo difference (stereo) channel, the Second Audio Program (SAP) channel, AM and FM radio stations, and radio subcarriers (SCAs), such as Radio Reading Services. Talking Books and described videotape are also mentioned in this section, though they are not broadcast technologies. The advanced technologies that will be discussed in Section 9.0 are use of the Vertical Blanking Interval (VBI), advanced speech synthesis over the closed captioning channel, synchronous audio tape (which is an issue for stations, not consumers), the Professional (Pro) channel, developing new TV audio channels, and advanced television (ATV).

The Stereo Sum (Main Audio) Channel

There is really only one issue associated with broadcasting video description over the main TV audio channel: Every viewer must listen to the description. A few small networks consider that to be an advantage because their programming is targeted at older and/or visually impaired viewers. However, major networks cannot assume that most of their viewers want video description with no option to turn it off. For major networks, the stereo sum channel is not a realistic option for video description.

The SAP Channel

The only major network that broadcasts described video is PBS, which relies primarily on the Second Audio Program (SAP) channel for video description. The SAP channel was chosen for three reasons:

Table 3. Technologies Capable of Broadcasting Described Video

Broadcast Technology	Pros	Cons
MTS Television Stereo System (Advantage: Sound is connected to picture.)		
Stereo Sum (Main Audio) Channel (15 kHz bandwidth)	All households can receive. All stations can transmit.	Reception not optional, so impractical for major networks.
Stereo Difference (Stereo) Channel (15 kHz bandwidth)	25% of households can receive, increasing. 48% of stations can transmit. Reception optional.	Only larger TVs can receive stereo now. Conflicts with stereo programs, and networks fear use would cause switching errors.
Second Audio Program (SAP) Channel (10 kHz bandwidth)	25% of households can receive, increasing. 20-48% of stations can transmit. Reception optional. 10% of stations use.	Only larger TVs can receive SAP now. Requires network to carry extra audio channel. Conflicts with second-language broadcasts, when available.
Professional (Pro) Channel (3.5 kHz bandwidth)	At most, 48% of stations can transmit. Reception optional. At most, 10% of stations use.	Practically no TVs can receive Pro now. Requires network to carry extra audio channel. Conflicts with intended use: station telemetry and cueing crews. Need signal processing to compensate for narrow bandwidth.
Special Modulation Techniques for TV (Advantage: Sound is connected to picture.)		
Vertical Blanking Interval (VBI) on TV Station (narrow bandwidth if only one VBI line used, easier with more than one line)	VBI can probably be routed through a major network's routing system and consoles without compromising on program timing.	VBI lines are in demand, but line(s) must be assigned to DVS. If VBI is used for final broadcast, need special receiver. Development required.
Advanced Speech Synthesis on Closed Captioning Channel	Narrow bandwidth required may permit sharing closed captioning VBI line without conflict. Sending pronunciation cues could make sound better than text-to-speech. Sharing closed captioning VBI line guarantees channel availability. Can probably be recorded on most VCRs.	Speech quality must be investigated. Special decoder needed, based on decoders that will be required for closed captioning starting in 1993. Regulation required.
SCA on TV Station (narrow bandwidth)	Not widely used.	Need special receiver. Development required. Probably not technically feasible on station already using all MTS channels. Requires network to carry extra audio channel.
Spread Spectrum on TV Station	Used successfully in U.K. for high-fidelity sound (NICAM system).	Need special receiver. Development required. Regulation required. Requires network to carry extra audio channel.

BEST COPY AVAILABLE

Table 3. Technologies Capable of Broadcasting DVS (Continued)

Broadcast Technology	Pros	Cons
Radio Modulation Techniques (Disadvantage: Sound is not connected to picture.)		
Main Channel of FM or AM Radio Station (15 kHz bandwidth or 5-10 kHz bandwidth)	Accessible virtually anywhere by anyone (for example, in cars). May attract general audience even without picture.	Air time is expensive. Simulcast requires network to carry extra audio channel or synchronize tape.
SCA on FM Radio Station (5 kHz bandwidth)	Slots increasingly available. Two SCA channels per radio station.	Need special receiver. SCAs may be in higher demand than SAP channel. Simulcast requires network to carry extra audio channel or synchronize tape.
Radio Reading Services (which are SCAs) (5 kHz bandwidth)	Print disabled have access.	Only print disabled have access. Limited number of Radio Reading stations. Simulcast requires network to carry extra audio channel or synchronize tape.

1) Every TV station has a SAP channel allocated to it that does not interfere with the main audio, even if the program is in stereo.

2) The SAP channel can be received by off-the-shelf stereo TVs and VCRs.

3) The SAP channel provides good sound quality (about 60 dB signal-to-noise ratio, @ 10 kHz bandwidth according to the Smith-Kettlewell report).

The first criterion alone (an audio channel allocated for every TV station that does not interfere with stereo broadcasts) excludes all but the SAP and Pro channels, but Pro receivers are not mainstream consumer products, and the sound quality Pro can provide is limited by its narrow bandwidth (3.5 kHz vs 10 kHz for SAP). Improving sound quality despite narrow bandwidth would add to the cost of Pro transmitters and receivers. Thus, from the consumer standpoint, SAP is the best way to transmit a second audio program to a television because that is exactly what it was designed to be.

As of early 1992, 61 PBS stations, in 23 states and the District of Columbia, carry some described video on the SAP channel. According to WGBH, together those stations cover over 46% of the 93 million households in the U.S. with a television, as shown in Table 4, from WGBH. That estimate does not take into account whether households within the viewing area of a station have a stereo TV or VCR to receive the SAP channel, but it does not double count households with access to several PBS stations. Now that PBS produces its programs in stereo, the number of PBS stations with SAP capability is rising because stereo-equipped stations can add a SAP generator more easily.

Table 4. Estimates of TV Households that Can Receive Descriptive Video Service

Station	Market Area	TV Households	Station	Market Area	TV Households
KAET-8	Phoenix, AZ	1,052,220+	KCPT-19	Kansas City, MO	
KUAS-27	Tucson, AZ		KETC-9	St. Louis, MO	1,088,550+
KUAT-6	Tucson, AZ		KTNE-13	Alliance, NE	
KCET-28	Los Angeles, CA	5,026,300+	KMNE-7	Bassett, NE	
KVIE-6	Sacramento, CA	1,033,780+	KHNE-29	Hastings, NE	
KPBS-15	San Diego, CA		KLNE-3	Lexington, NE	
KQED-9	San Francisco, CA	2,231,040+	KUON-12	Lincoln, NE	256,900+
KRMA-6	Denver, CO	1,053,510+	KRNE-12	Merriman, NE	
WETA-26	Washington, DC	1,749,190+	KXNE-19	Norfolk, NE	
WLRN-17	Miami, FL	1,313,540+	KPNE-9	North Platte, NE	14,380+
WSRE-23	Pensacola, FL	417,600+	KYNE-26	Omaha, NE	347,160+
WEDU-3	Tampa, FL		WNET-13	New York, NY	7,043,790+
WTTW-11	Chicago, IL	3,141,510+	WMHT-17	Schenectady, NY	491,980+
WNIT-34	South Bend/Elkhart, IN		WCNY-24	Syracuse, NY	370,870+
KDIN-11	Des Moines, IA	373,060+	WUNC-4	Chapel Hill, NC	
KIIN-12	Iowa City, IA	321,080+	WEAO-49	Akron, OH	277,130+
WCBB-10	Lewiston, ME	345,370+	WNEO-45	Alliance, OH	
WMPT-22	Annapolis, MD		WBGU-27	Bowling Green, OH	52,030+
WMPB-67	Baltimore, MD	938,520+	WCET-48	Cincinnati, OH	766,730+
WFPT-62	Frederick, MD		WVIZ-25	Cleveland, OH	1,460,020+
WWPB-31	Hagerstown, MD	43,300+	WGTE-30	Toledo, OH	414,230+
WCPT-36	Oakland, MD		KOAB-3	Bend, OR	27,200+
WCPB-28	Salisbury, MD	95,040+	WLVT-39	Allentown, PA	
WGBH-2	Boston, MA	2,141,400+	KLRU-18	Austin, TX	
WGBY-57	Springfield, MA	218,990+	KERA-13	Dallas, TX	1,735,380+
WUCX-35	Bad Axe/Ubly, MI		KUHT-8	Houston, TX	1,471,840+
WTVS-56	Detroit, MI	1,722,470+	KUED-7/9	Salt Lake City, UT	592,200+
WFUM-28	Flint, MI	454,130+	KCTS-9	Seattle, WA	1,321,920+
WGVU-35	Grand Rapids, MI	626,440+	WSWP-9	Beckley, WV	144,720+
WGVK-52	Kalamazoo, MI		WMVS-10	Milwaukee, WI	772,710+
WUCM-19	Univ. Center, MI				

NOTE: The above list includes 9 of the top 10 markets in the U.S. [Jan 91 Nielsen U.S. TV Est.]

Total TV households capable of receiving DVS = 42,948,230 million
Total TV households in the United States = 93,046,390 million

Stereo TVs and VCRs can almost always receive the SAP channel, according to WGBH. The Electronic Industries Association provides the following statistics on stereo TVs:

- In mid-1990, 98% of households had at least one TV, 96% had a color TV, 56% had a monochrome TV, and 21% had a stereo color TV. Most households have more than one TV.

- Stereo TV sales doubled from 1985 to 1986, and had doubled again by 1990.

- Since the introduction of MTS stereo in 1984, an estimated 34 million stereo color TVs have been sold, out of the 157 million color TVs sold in that time (22%). About 29% of the sets sold from 1988 to 1991 are stereo TVs. Of the estimated 22 million color TVs sold in 1991, over 7 million were stereo TVs (32%), accounting for much of the growth in color TV sales.

- Sales of stereo color TVs have exceeded sales of monochrome monaural TVs, including those sold for use with computers, since 1987. In 1991, stereo color TVs outsold monochrome monaural TVs more than 5 to 1.

The Electronic Industries Association provides the following statistics on stereo videocassette recorders (VCRs):

- In mid-1990, about 69% of U.S. households had at least one VCR.
- Stereo VCR sales, in terms of number of VCRs sold, were growing at a rate of about 23% in 1990, despite the downward trend in overall VCR sales.
- Of the estimated 9.5 million VCRs sold in 1991, about 2 million were stereo VCRs (21%).
- Since 1988, 6.7 million stereo VCRs have been sold, out of 40 million VCRs sold in that time (17%). That percentage is growing, because non-stereo VCR sales are slowly falling while stereo VCR sales are slowly growing.

These statistics mean that over 21% of households can receive the SAP channel on a stereo TV. Sixty-nine percent of households have VCRs, and on the order of 10% of those VCRs are stereo VCRs. Many people willing to spend the extra money for a stereo VCR may also have a stereo TV, so stereo VCRs can only increase the number of SAP-capable households by a few percent over the estimate based on stereo TVs alone. For a round estimate, that means roughly 1 in 4 households (25%) are SAP-capable in early 1992, with that percentage increasing rapidly as more households replace non-stereo TVs and VCRs with stereo-equipped models.

Almost half of Americans live in an area where described video is broadcast on PBS, and that percentage is also rising. Together, those figures do not indicate what percentage of households are both situated and equipped to receive described video on PBS beyond setting an upper limit of around 25%, but PBS stations are generally concentrated near cities and metropolitan areas. Since stereo TV stations are similarly concentrated, presumably so are stereo TVs and VCRs. That would mean the percentage of households capable of actually listening to a described video broadcast today may be as high as 20%--15 or 20 million households, but a survey would probably be necessary to answer the question with any certainty.

The percentage of the visually impaired population that is both equipped and situated to receive described video on PBS is even more difficult to estimate. Concentration near cities may mean the visually impaired are more likely to live near a SAP-capable PBS station, but the finances of people with disabilities and the reduced incentive to own a more costly TV or VCR (given visual impairment and little described video now available) probably make owning an appropriate receiver less likely for the visually impaired. The Smith-Kettlewell study found that there is no reason why only color TVs with a 20-inch-diagonal picture or larger should be available with SAP. All that is required is that "the audio bandwidth of the detector [be] broad enough." The lack of lower-priced TVs and TV radios with SAP capability may not be as glaring an issue to those with milder visual impairments, but blind people would obviously benefit from SAP capability on smaller TVs and TV radios. Aside from entertaining visitors, a person with little or no residual vision, living alone, has little or no use for the parts of a television that account for most of its cost, power consumption, and limited portability: the picture tube and associated electronics. The vast majority of people with severe visual impairments have at least some residual vision, though. One thing is clear: people with all levels of visual impairment would have more incentive to get SAP-capable receivers if more programs were described on the SAP channel.

A more subtle but important issue is whether consumers, especially those with visual impairments, know how to access the SAP channel on their TVs and VCRs. Remote controls are notoriously difficult to figure out how to use, instructions are notoriously poor, and a visually impaired person who has a stereo TV or VCR will not magically know how to access the SAP channel with it. Even with large-print or Braille instructions, that is a serious problem, especially since few TV or VCR manufacturers would think of the visually impaired as potential customers. Probably few consumers have even heard of the SAP channel, and fewer know how to access it. That problem must be solved for there to be any chance of making DV on the SAP channel very useful, let alone commercially viable. Consultation with equipment manufacturers and an education campaign are the obvious solution to this problem.

PBS sends finished programs to its broadcasting stations via satellite, either for live broadcast or for recording and later transmission. The narration is sent over a separate satellite audio subcarrier using standard equipment. Then it is patched to the SAP generator at the transmitter.

With "American Playhouse," in 1988, WGBH and PBS demonstrated that satellites can carry the audio subcarrier for video description from the TV network to individual stations over a six-month test period. The problem is, commercial stations have different priorities because they have to make a profit; there is no guarantee that they will follow, producing and broadcasting described video.

As of late 1991, the SAP channel is underutilized, often because stations lack a SAP generator and/or the financial incentive to use one. For example, a simple informal survey, by SAIC's staff, of the channels available from a basic cable system in the Washington, D.C. metropolitan area (repeated on several different days and at various times of day) showed the SAP channel was used by two of the three local PBS stations and by Home Box Office (HBO). SAP transmission was detectable on all three channels from about 7 p.m. to midnight. One of the PBS stations uses its SAP channel to transmit weather, regardless of what is on the main audio channel. HBO was using its SAP channel for Spanish-language translations of what is on the main audio channel. When translations are not available, they turn their SAP generator off. Similarly, the second PBS station uses SAP for video description or Spanish-language translations, whichever is available. Other common uses for SAP are duplicating a radio station's programming, broadcasting music, or reading program listings. The COSMOS report gave the estimate that 10.3% of network-affiliated stations use the SAP channel; 48% have MTS stereo capability. Roughly 20% of stations have a SAP generator.

AM and FM Radio Stations

AM and FM radio is the most popular broadcast communications medium in the U.S. According to the Electronic Industries Association (EIA), 98% of households have at least one radio, and radio sales exceed TV sales. The market for receivers is huge and mature, so almost any conceivable feature is available and affordable. Many visually impaired people listen to the radio because radios provide convenient portable entertainment and news at a bargain price. In fact, before television, radio shows were much like described television shows, ideal for visually impaired people, and sporting events on the radio are still much like the old radio shows. It may not be possible or desirable to bring back the days of the old radio shows, but there is likely to be a market for that kind of show, not only for the visually impaired but for commuters, truckers, factory workers, etc. Video description produces material suitable for that type of audience, and it would not always have to be simulcast (simultaneously broadcast) with relevant video. By simulcasting the described audio program through radio routing, special handling procedures to route the described video track onto the TV broadcast would not be necessary. However, if the video description track on the video tape is simulcast on the radio, this track is physically a special row (stripe) on the broadcast tape, so a technician would have to patch this new track into the distribution system at the broadcast station. If the technician forgot to unpatch that track, the station would experience dead air (silence) or broadcast audio time code instead of the intended sound track.

Subsidiary Communications Authorizations (SCAs)

Fourteen FM radio stations in the U.S. now simulcast PBS described video over Subsidiary Communications Authorizations (SCAs), at least some of which are Radio Reading Services. Table 5, from WGBH, lists the 14 stations. An ordinary FM radio cannot receive these broadcasts. Radio Reading Services have been around since 1969, and in 1983 there were 85 of them in the U.S. and Canada [*That All May Read*, Library of Congress National Library Service for the Blind and Physically Handicapped, 1983]. Visually impaired listeners access Radio Reading Services with a special pretuned receiver, available free to anyone who is unable to read print.

In technical terms, SCAs are FM subcarriers at either 67 kilohertz or 92 kilohertz. Since two subcarrier frequencies have been authorized by the FCC, every FM station can carry up to two SCA channels at once. SCAs have been used for various purposes, but supply now exceeds demand because their use for broadcasting background music to places like shopping malls has been displaced by the use of satellite links. On the other hand,

Table 5. PBS Stations that Simulcast Descriptive Video Service (over FM subcarriers or Radio Reading Services)

Station	Location	Names of Subcarrier
KPBS	San Diego, California	Radio Reading Service
WLRN	Miami, Florida	Radio Reading Service*
WILL	Urbana, Illinois	Illinois Radio Reader
WTVP	Peoria, Illinois	WCBU Radio Service
WNIN	Evansville, Indiana	Radio Reading Service
WGBX	Boston, Massachusetts	TIC Radio*
WMHT	Schenectady, New York	RISE Service
WCNY	Syracuse, New York	Read-Out
WXXI	Rochester, New York	Reachout Radio
WVIZ	Cleveland, Ohio	Radio Reading Service
KOAC	Corvallis, Oregon	Golden Hours
KTVR	LaGrande, Oregon	Golden Hours
KOAP	Portland, Oregon	Golden Hours
WHRO	Norfolk, Virginia	Hampton Roads Voice

*Also carries DVS on TV SAP channel.

Smith-Kettlewell found they are increasingly being used for computer data transmission. There are several major disadvantages to using SCAs for video description. First, Radio Reading Service receivers are pretuned to a single station, limiting the number of potential simultaneous video descriptions broadcast. Adding more stations would require adding new receivers or designing and distributing multifrequency receivers, which could prevent their becoming very popular. Even with a special receiver, finding what TV picture is associated with what audio description channel would be a nuisance for people who are less severely visually impaired, and the chance for finding a broader market would be lost. With the SAP channel, the association between the picture and audio description is automatic, universal and intuitive. Video description over Radio Reading Services could also divert the use of Radio Reading Services from their original purpose: reading books and sharing other information for the visually impaired. For an occasional broadcast, that diversion is not an issue, but for widespread use, Radio Reading Services cannot be casually pushed aside. These functions are at least as important as described video to many people. Smith-Kettlewell points out another important disadvantage. Radio Reading Services have a narrow bandwidth (5 kHz, comparable to AM radio), and are subject to crosstalk--distorted sound from the radio station's stereo channel. For reading a book, that may be acceptable, but the sounds of a television program combined with video description are more complex, and the reduced sound quality would be quite significant. Thus, Smith-Kettlewell recommends Radio Reading Services as a backup for areas that do not presently have SAP capability for DV.

Talking Books

Talking books have been around for over 50 years, giving the visually impaired access to books read by a narrator. Talking books, like Radio Reading Service receivers, are available free to anyone who is unable to read print. They could also be used for the narrated sound track to movies, for example, but they would not allow less severely visually impaired people to watch the movie at the same time, synchronized with the sound track. Narration broadcast simultaneously with video (simulcast description) does that automatically. Also, simulcast DV does not have to be ordered by the viewer. There is no turn-around time. In short, talking books are a valid medium for described video for the blind, but they are much less flexible than broadcast described video and are not a substitute. It would also be inefficient to go to the expense of producing described video that can be synchronized with the corresponding video unless the video were also made available with the described audio. Again, Smith-Kettlewell recommends talking books as a backup for areas that cannot yet receive broadcast described video.

Described Videotape

Video description on videotape offers several advantages over the use of talking books. First, the visually impaired have access to the video as well. Second, if it is desirable to put the description on the main audio channel, that is also possible, if the videotape is targeted at described video users. Hi-fi stereo VCRs can actually record and play three audio tracks on a videotape besides the video, allowing the stereo sound to be recorded in hi-fi stereo with the (monophonic) SAP channel on the third audio track. However, only two of the audio tracks can be played back at a time, and hi-fi stereo VCRs do not generally allow mixing their monophonic audio track to be played back with one or both of the stereo tracks. Modifications to the hi-fi stereo VCR would have to be made in order to be able to playback the third audio track mixed with the stereo channels, though this would only be necessary if narration were transmitted without mixing with the main sound track. VCRs that cannot play videotapes in stereo would only play the videotapes recorded with stereo audio plus monophonic SAP. The format of videotapes with video description should be given careful consideration, because videotape does not provide nearly as many options as the MTS Stereo system provides. For example, it is not necessarily true that recording a broadcast with video description on the vertical blanking interval (VBI), which will be described later in this scenario, will allow the description to be retrieved on playback with a decoder. That might be the case with studio VCRs but it may not be true of home VCRs. Testing is required, and is presently underway via a grant to WGBH from the Department of Education's National Institute on Disability and Rehabilitation Research (NIDRR); closed captioning has a relatively low data rate, so it is not necessarily indicative of what would happen with higher data rates.

6.0 THE DEPARTMENT OF EDUCATION'S PRESENT COMMITMENT AND INVESTMENT

- The Department of Education, Office of Special Education Programs (OSEP) is currently funding two descriptive video projects, both at station WGBH, Boston. One program pertains to descriptive video in the Public Broadcasting System (PBS), the other is for descriptive video on videotape.
- The Department of Education's NIDRR is currently funding a WGBH study to evaluate the use of the vertical blanking interval in television broadcasts to include descriptive video in commercial television broadcasts.
- The Panel of Experts for the Department of Education program sponsoring this study consists of experts from industry and Government, including members of the sensory-impaired community. Their consensus opinion was that developing descriptive video is a priority for persons with visual impairments.
- One of the Department of Education's 1991 Small Business Innovative Research (SBIR) Program Research Topics is to develop or adapt communication devices for young children who are blind. Descriptive video could be used to enhance educational television programming to provide equal access for youths with vision impairments.

Two of the Department of Education's 1992 Small Business Innovative Research (SBIR) Program Research topics directly apply to DV:

- "Adaptation or development of devices to provide individuals who are blind with closed audio track to explicate the visual, non-verbal, non-auditory features of television and movies."
- "Exploration of alternative technologies for providing Descriptive Video Services (DVS) to persons with visual impairments."

7.0 ACCESS TO COMMUNICATIONS MEDIA

Many federal, state, and local laws influence access for persons with visual impairments. The most important single law related to access for persons who are vision impaired is Public Law 101-336, enacted July 26, 1990. Better known as the Americans with Disabilities Act (ADA), this law has broad implications for all disabled Americans and establishes the objective of providing access to persons with disabilities to physical and electronic facilities and media.

The other law that impacts technology for persons with visual impairments is Public Law 100-407-AUG. 19, 1988 titled "Technology-Related Assistance for Individuals with

Disabilities Act of 1988." Also known as the Tech Act, this law established a comprehensive program to provide for technology access to persons with disabilities. The law defines assistive technology devices:

> "Assistive technology devices means any item, piece of equipment, or product system, whether acquired commercially off the shelf, modified, or customized, that is used to increase, maintain, or improve functional capabilities of individuals with disabilities."

Descriptive video clearly meets this definition for persons with vision impairments and should be exploited to increase the ability of persons with vision impairments to obtain access to the medium of television. Within the findings and purpose of this law, descriptive video can provide persons with vision impairments with opportunities to:

- exert greater control over their own lives by making television viewing more realistic and understandable;
- participate in and contribute more fully to activities in their home, school, and work environments, and in their communities;
- interact with non-disabled individuals by providing the ability to talk about the programs they listen to on television; and
- otherwise benefit from opportunities that are taken for granted by individuals who do not have vision disabilities.

The Government regulations that affect the technical aspects of implementing DVS come from the Federal Communications Commission (FCC). The FCC has regulated the most promising modes of delivering DV so that they may be legally used for DV and for many other services. That is good for DV users, because the best channels are available; but, as the FCC intended, there is always competition for them.

The SAP channel may also be used for second-language broadcasts (typically, Spanish), 24-hour weather, music, or anything else a TV station chooses to broadcast (and monitor). Most stations do not use the SAP channel at all. For the time being, the flexible approach taken by the FCC generally works well; the SAP channel is shared among services at the stations' discretion. Some stations dedicate their SAP channel to a 24-hour service, but most of those stations would probably not use their SAP channel at all otherwise. The problem comes in when services expand to the point where they frequently preempt other services. Then, there may be a need to reserve channels for specific services, potentially compromising on the choice of transmission channel for one that will never be preempted. Of course, dedicating an audio channel, such as the SAP channel, to any one service closes it to all other services, which may result in inefficient use of the resource.

As with the SAP channel, there is no law against using one or two lines of the vertical blanking interval (VBI, the invisible lines at the top of a TV picture), for DV transmission, as discussed in section 9.0 of this scenario. However, regulation would be necessary to standardize which line(s) and reserve them for DV. Otherwise, different stations would use different lines, making reception by the consumer impractical, and there would be direct competition between DV and teletext. Since teletext is basically transmission of text (and pictures) to paying subscribers, DV would almost certainly not be as profitable and never establish itself without FCC regulation to dedicate a line or two of the VBI to it. This was done for closed captioning in 1980, where all of field one and half of field two of line 21 was authorized by the FCC.

In short, the regulations affecting DV implementation permit it but do not guarantee that it will be provided or that it will not be preempted by other services.

8.0 POTENTIAL ACCESS IMPROVEMENTS WITH ADVANCED DESCRIPTIVE VIDEO TECHNOLOGY

Many recent technological advances make DV possible. The proliferation of small computers has made creating and editing narration far less tedious and expensive. Satellites have made it much less expensive to send the narration from network to affiliates, and stereo TVs and VCRs have brought the ability to receive the SAP channel into over 20 million living rooms. Other transmission options have been opened by advanced signal processing and audio data compression technology, making it possible to send more sound in less bandwidth, over channels that would otherwise be too narrow. The use of time code in the broadcasting industry has made it possible to synchronize narration with a TV program, even if the program is later edited. The advent of MTS stereo in the mid-80's has made the SAP channel available to transmit the narration. Finally, if the vertical blanking interval of the television picture were used to transmit DV digitally, new pulse code modulation and linear predictive coding (PCM and LPC) chips could reduce the size and cost, and raise the performance of the decoders. Another possibility would be the use of audio compression technology, which can reduce the bandwidth requirements of compact-disc-quality audio by a factor of four without any perceptible loss in sound quality.

9.0 ADVANCED DESCRIPTIVE VIDEO TRANSMISSION TECHNOLOGIES

The obstacles to putting video description on the SAP channels of commercial stations are essentially cost and availability issues. The actual production costs of video description are reduced by technology, but technology cannot eliminate them. They are covered in the next section. However, many of the distribution problems associated with video description can be solved or minimized through advanced technology, including competition for the SAP channel between DV and second-language broadcasts.

Routing An Extra Audio Channel Through A Television Network Facility

Commercial network facilities were simply not designed to handle three channels of audio, and Smith-Kettlewell found that to be the most important reason why networks are afraid to try broadcasting described video. Networks frequently broadcast programs in stereo, so two channels are not a problem, but mainstream network control equipment was not designed to work with a third audio channel. That leaves four options open:

1) *The networks can buy special routing switches and rewire extensively, solving the problem directly* (and upgrading the networks' equipment), but at an estimated cost, according to the Smith-Kettlewell report, of 10-20 million dollars, much of that being labor costs. Eventually, that sort of upgrade will occur anyway, but there seems to be no reason to believe it will be soon.

2) *The networks can handle the third channel separately.* For cost reasons, that would involve some manual switching, called "patching," instead of the relatively recently established norm of automated switching. They could patch the third audio channel around the audio console and directly into the SAP generator (exciter). The networks fear this would cause human error and imprecision in the switching, acceptable for what Smith-Kettlewell refers to as "special events," but problematic for day-to-day operations. In essence, treating DV as a special event would limit its use to one or two programs a week, because it would require special network accommodations to be made every time a described program is broadcast, including the use of extra audio consoles and temporary cabling. There would also be a start-up cost of about a million dollars, according to the Smith-Kettlewell report.

   An alternative would be to use a low-power (less than 1-watt) radio transmission system within the network and/or station facility. This system could transmit the third audio channel, switched in by a coded signal, thus bypassing the audio console. This approach could save some of the cost and inconvenience of temporary wiring in some cases, providing an interim solution until the network or station upgrades its system. Innovation is key to this approach. Essentially off-the-shelf equipment could be used.

3) *The networks can consider video description as the second channel, and never broadcast described programs in stereo.* Smith-Kettlewell found that PBS considers this practical, but the commercial networks are afraid it would cause confusion and lead to human error in arranging the automatic patching.

The Vertical Blanking Interval (VBI)

A broadcast television picture in the U.S. (unlike a computer display) consists of 525 lines of video, transmitted 30 times per second. A technique called "interlacing" is used, whereby every other line (half of the lines) is transmitted; then the lines in between (the other half) are transmitted. Each half is called a field, so 60 fields are transmitted every second. It takes a small fraction of a second for the electron beam, which paints the television picture, to get from the bottom line to the top line of a television screen. During that time, the electron beam must be turned off, or "blanked"; otherwise, a bright diagonal line would appear on the screen. The time that the beam is turned off is called the vertical blanking interval, or VBI.

During each VBI, about 21 lines could have been transmitted, but the electron beam is off ("blanked"), so that its rapid diagonal motion does not show up on the screen. It is possible, however, to use that time interval to send data without interfering with the television picture. For example, line 21 of the VBI is reserved for the text of closed captioning for the hearing impaired. Lines 1-9 are reserved to ensure that the picture does not roll. According to Kelly Williams, of the National Association of Broadcasters, using lines 10-14 can cause bright dots to appear on TVs sold before 1975. Field 1 and half of field 2 of line 21 are reserved for closed captioning, though only field 1 is used so far, and line 22 appears at the top of many newer TV screens. That leaves lines 15 through 20 available. Lines 10-14 may also be worth considering, as older TVs go out of use. Six to eleven lines would be far more than is needed for an extra audio track, but stations also use their VBI lines as a source of income. VBI lines can be used to transmit text and images or for other transmission services. PBS even used its VBI to transmit digital audio to its affiliate stations, via satellite, in the early- to mid-80's. According to the Smith-Kettlewell report, "Video tape recorder manufacturers are considering this idea as well; more 'audio tracks'...would make room for special-effect channels, etc."

Teletext is relevant to DV for three reasons: it may compete with DV for VBI lines, it gives a feel for the data rates that can be achieved over VBI lines, and it may offer a format to use for transmitting digital audio. After all, digital data is digital data whether it is digital audio, text, or pictures.

There are two systems used for teletext in the U.S.: CBS uses the North American Basic Teletext Specification (NABTS), and Taft Broadcasting uses World System Teletext (WST). For teletext, the two formats are incompatible, but their digital data format appears to be the same. Both use a data rate of 5,727,272 bits per second (a bit is a binary digit: 0 or 1). That is a huge data rate for sound or text, but it must be divided by the number of lines in the picture, and there is some overhead, including synchronization and error detection. The actual data rate of both systems turns out to be 11,760 bits per second per VBI line, so, considering that error detection is not error correction, one line would convey intelligible speech. More than one line would probably attract a wider audience, especially if the current practice of mixing the narration with the main sound track is followed.

EEG Enterprises makes closed captioning equipment, but they also make VBI transmission equipment. They can send 9600 bits per second per VBI line, "virtually error-free." At 4800 bits per second per VBI line, the information on the VBI can be recorded on a non-studio VCR.

The great advantage of putting audio information on the VBI is that a television network can work with it implicitly. Wherever the video goes, the VBI automatically goes with it; there is no need to spend millions of dollars to replace network routing equipment and consoles. It is also possible, given enough VBI lines to work with, for networks to offer more than one alternative audio program on the VBI. If local stations rebroadcast the audio from the VBI on the SAP channel, that would let the local stations determine which to rebroadcast based on local requirements. For example, Spanish translations of programs would certainly be in demand in southern California, but video description might be in greater demand in Boston. Perhaps stations in Florida would broadcast programs both ways, depending on the time of day, or simulcast popular described programs on a radio station.

Fifty dollars or less would be a rough estimate of the cost of a decoder for DV on the VBI, given the maturity of the equipment and given that closed captioning decoder chip sets costing $5 could be modified for this purpose. International Telephone and Telegraph (ITT) makes one such chip set. New technologies that could be used for such equipment would include pulse code modulation and linear predictive coding (PCM and LPC) chips to reduce the size and cost of the decoders, audio compression technology, which can reduce the bandwidth requirements of compact-disc-quality audio by a factor of four without any perceptible loss in sound quality, and neural networks, which might be able to help with correcting transmission errors.

There are very strong arguments for decoding the DV audio at the transmitting station and broadcasting it on the SAP channel as well as, or instead of, on the VBI. Transmitting DV into homes on the VBI could make it underutilized, due to the need for a special DV decoder. When a fixed-cost service for the visually impaired can serve everyone, there is no reason not to take full advantage of the service. In this case, the cost of describing a program does not increase if ten times as many people watch it, and using a decoder that is considered special equipment for the visually impaired might create the same problem experienced with closed captioning and access to text. The cost of a VBI decoder at a station's transmitter facility was estimated at less than $200, and there would be no need to use an extra satellite link, microwave link, or anything else; just a SAP generator and MTS-capable transmitter.

PBS may also adopt this approach, since they encounter many of the same cost and convenience issues faced by other networks in upgrading to SAP capability. VBI should be fully explored for DV.

Speech Synthesis for the Narration

If narration is transmitted in digital form over one or more VBI lines, the digital format can be optimized to get the best sound quality possible for a given number of VBI lines. Compact disc data rates would require on the order of 70 video lines, minimum, and advanced digital audio technology can get the same audio quality at a quarter of that data rate. However, even 15-20 video lines would require a TV station that has no picture. The only reason for using a TV station with no picture would be to get the best sound quality possible, but it would be much more desirable to keep the sound together with the picture. Video description would seldom benefit from compact-disc sound quality anyway, and a TV station dedicated to DV and teletext may be cost-prohibitive. It is unlikely that more than a few VBI lines of a TV station with a picture would be dedicated to DV in the near future because there are only about 5 or 10 VBI lines per station that can be used to transmit data. Without dedicated lines, DV would have to compete with profit-generating services; essentially, variations on teletext.

All this leads to one conclusion: some sound quality must be sacrificed to minimize the number of VBI lines required by DV. With a few VBI lines, sound quality would probably be acceptable. With one VBI line, intelligibility may be an issue. Half a VBI line would probably not be feasible.

There are ways to transmit speech over half a VBI line though, if it is not important that the voice heard at the receiving end sound like the voice at the broadcasting end. In its simplest form, the narration could be sent as text, read by a speech synthesizer in the users' home. This technique could result in data rates as low as 300 bits per second (bps), requiring only a fraction of a VBI line. At such low data rates, it would be technologically feasible for video description to share the VBI line used for closed captioning (i.e., the text channel). Sharing closed captioning's VBI line should not create any conflicts, because closed captioning only uses part of the line allocated to it anyway.

Smith-Kettlewell estimated the cost of text-reading DV receivers at almost $2000, but that assumes a costly proprietary speech synthesizer would be needed for high quality synthetic speech from text. If that estimate is valid, it would raise questions about whether the cost of producing the narration for synthesis could be justified by the number of users it would attract, unless video description is also available through other channels. However, to a machine, pronouncing text is much more difficult than pronouncing phonetic information, and it is probably a more efficient application of technology to give a much less expensive speech synthesizer more phonetic information than just text. The most difficult (and expensive) part of developing a machine to pronounce text is getting it to sound more human and pronounce words correctly, accenting the right syllables, raising and lowering the pitch and loudness of its "voice," and pacing itself correctly. Therefore, it makes sense to consider giving the machine something easier to pronounce than text.

A person who is severely visually impaired may own a speech synthesizer already, suggesting still another approach. If text is what is transmitted, people could then use their own speech synthesizers to read it, given computer programs and electronics to make that possible. That approach may not be feasible, however, for three reasons. First, the cost of developing and maintaining the software and hardware to accommodate different speech synthesizers may be prohibitive. Second, although people may be able to understand speech from their own speech synthesizers best, they might not want to listen to the same speech synthesizer all day long, both for work and for entertainment. Finally, there are many visually impaired people who do not own speech synthesizers, and this approach would not help them at all. It is also extremely unlikely that anyone who is not severely visually impaired would consider using such a service, if for no other reason than the cost of high-quality text-to-speech equipment and the time it takes to learn to understand even high-quality synthetic speech from text.

For their report, Smith-Kettlewell conducted a simple experiment and found that 2 out of 10 members of their panel considered the speech of even the $3,500 DecTalk (Version 2.0) speech synthesizer "objectionable enough not to recommend it for DV." Their panel was composed of people who are used to listening to synthetic speech: visually impaired people who use talking computers. Based on that simple experiment, text-to-speech technology may not be feasible for transmitting DV now, but that could easily change in a few years. A compromise between the low bandwidth requirement of text transmission and the better sound quality of digital audio transmission is already possible. Speech synthesizers such as the compact RC Systems V8600/1, available for $99 in quantity, may be able to bridge the gap, given pronunciation information instead of just text.

Most of the cost of producing DV is in composing the script. Thus, cost savings at the DVS production end may not justify the loss of audience that speech synthesis might bring. The real cost savings would be in the ability to deliver video description over a single VBI line, perhaps even sharing the line with closed captioning. If only a handful of people use synthesized DV because the sound quality is low and receivers are expensive, it probably makes more sense to invest in putting DV on the SAP channel. That way, the costs are higher, but the benefits would go to a much larger and more diverse audience. Low-quality synthetic speech from text makes little sense as a backup approach to delivering DV, because a person whose vision is impaired to the point where he needs a high-quality speech synthesizer would probably be best served by a SAP-capable TV radio. Customized VBI decoders for specific speech synthesizers would probably have almost no market, so they would cost more than SAP-capable radios, and they would not even approach the sound quality and larger-market conveniences offered by the SAP channel. However, synthesized DV may be appropriate if decoders can be incorporated into TVs using the chip sets for closed captioning, required as of 1993. The possibility of transmitting DV in the form of pronunciation data should be investigated further.

Synchronous Audio Tape

Smith-Kettlewell suggests another way to get around the problem of sending three channels of audio through a network facility wired for only two channels. Bypass the network facility, and originate the third audio channel elsewhere. The tricky part is to keep the audio in synchronization with the video, since a relatively short lag is perceptible, and short lags are sure to become long lags over the course of a program. The synchronization can be done with Society of Motion Picture and Television Engineers (SMPTE) time code. The network or station broadcasting the video conveys the time code from the video to a synchronous tape recorder which has the extra audio channel. The synchronous tape recorder, costing about $6,000, can then put out the audio in synchronization with the video.

Two tracks must be transferred for that to work. The time code has to get from the network to the synchronous tape recorder, and the audio from the tape recorder has to be broadcast. The synchronous tape recorder could be located either at a single site with a link to the network's satellite (for example, the network facility), or at every affiliate station carrying DV.

There are several problems with this approach, however. As Smith-Kettlewell points out, remote facilities require a staff and equipment, and that costs money. Also, the network still has to send time code to the facility or facilities that have the synchronous tape recorder. Also, "last-minute changes to the program (such as the insertion of announcements) would alter the length of the show, in which case the synchronous audio tape at the DV Facility would no longer match the program." It would take extra coordination to prevent that problem, although Smith-Kettlewell points out that there is always the possibility of giving someone a video monitor and a microphone with which to narrate live for a while, should the synchronous tape option experience temporary problems. Spontaneous narration, though theoretically possible, is not considered to be a feasible solution to tape problems, because it requires a competent describer to be available at all times and still produces narration that is greatly inferior to pre-planned narration.

In short, the use of synchronous tape recorders is a technically feasible option for handling the extra audio track for DV, but it is not clear that it would be any more practical, overall, than sending the audio over the VBI. Sending the audio over the VBI certainly seems more elegant, direct, and less prone to error, but that remains to be seen as the results of experiments with VBI become available through the WGBH/Department of Education grant late in 1992.

Smith-Kettlewell also suggests the possibility of a service, to which affiliate stations would subscribe, that transmits the narration track and an "Audio Time Code" track on separate satellite audio channels. That transmission could be done on off hours. Stations would record that information, in the form of a synchronous audio tape, for later transmission. Coordination problems would still be significant using this approach.

The Pro Channel

The Pro channel was originally intended to be used to cue station employees, telling reporters and camera crews when they will be on the air, and for station telemetry. Thus, it is not surprising that the Pro channel is not accessible to the overwhelming majority of stereo TVs and VCRs. Some stations may object to widespread use of Pro channel receivers, because Pro was not intended for reception by the public, and some stations may not want to give up their Pro channel, but those are not the main issues with using Pro. The real problems with Pro are that it is a narrow channel, with a bandwidth of only 3.5 kHz; it has no built-in noise reduction scheme, though one could be added; and consumer receivers for the Pro channel are practically non-existent. According to the Smith-Kettlewell report, "it is rather noisy in poor reception areas, and it is very badly affected by multipath distortion." The bandwidth of the Pro channel would certainly be adequate to carry the text of video description for synthesis, to be read by a speech synthesizer, but that would raise essentially the same problems as are associated with dedicating a VBI line to text for speech synthesis plus Pro's own problems. Transmitting audio over the Pro channel would require considerable signal processing to reduce noise, but the narrow bandwidth would still make its sound quality significantly lower than that of the SAP channel. Pro generators, with added noise reduction equipment, would probably cost thousands of dollars, as do SAP generators. Overall, the Pro channel may make a better backup channel than a primary channel.

Developing New TV Audio Transmission Channels

In television audio, all but the stereo sum (main audio) channel of the MTS system are subcarriers of the stereo sum channel. It is unlikely that more subcarriers (SCAs) could be added to the television audio signal unless a station is not using its stereo difference, SAP, or Pro channel. Too many stations broadcast in stereo to make using that space viable, and there is no incentive to add a non-standard channel only to give up the SAP or Pro channel. The problem associated with putting too many subcarriers in too little bandwidth would be interference such as "birdies," which are unacceptable chirping sounds, or worse, distorted voices from other MTS channels. The MTS system was designed to be compatible with monophonic TVs and easy to implement, minimizing noise, interference, and the effects of receiver location. Trying to improve upon the MTS system design by adding channels that are practical to use would be a major undertaking. It is legal for any TV station to add subcarriers that are monitored by the station and do not interfere with receivers following the MTS system. However, without a uniform standard, there is little chance those isolated systems will add up to a national DV capability, due to the cost and inconvenience of non-standard receivers, and the customer base for described programs would remain small.

One contact suggested spread spectrum modulation as a possible way to broadcast DV over TV stations without interfering with other broadcasts. Spread spectrum techniques were reportedly used in the British NICAM system for hi-fi TV audio broadcasts. A U.S. version would have to be developed, however, and then approved by the FCC. Since such a system would require special receivers to be developed and distributed, and significant interference with other broadcasts could not be tolerated, that development would be relatively risky, technically. Spread spectrum modulation should be investigated, but only to provide a backup solution if problems arise with using the VBI.

Advanced Television (ATV) and High Definition Television (HDTV)

As TV technology advanced over the years, screens have become larger, but the level of detail visible in them has remained essentially the same since the emergence of color television. Adding detail to a TV picture requires new standards: standards that will be adopted by the FCC by late 1993. Many approaches are being considered, but the eventual goal is High Definition Television (HDTV), possibly preceded by an intermediate step, Advanced Television (ATV). In 1990, the EIA estimated that it will be "at least 1993 or 1994 before consumers will be able to go into a store and order a new wide-screen receiver," but a more recent estimate would probably delay that by a year or two.

It will be relatively easy, from a technical standpoint, to incorporate a channel reserved for video description into ATV and HDTV. The important point is that the FCC must include a DV channel in their specifications or the present situation may persist for a long time. Retrofitting is always more difficult and expensive. Table 6 shows WGBHs suggested allocation for DV, based on a Cable Labs proposal.

Table 6. Advanced Audio Simulcast ATV

	Service	**Data Rate kbit/s**
PROGRAM RELATED SERVICE	Main Program, Four Channels	512
	SAP Stereo, Two Channels	256
	DV, One Channel (optional)	128
	Expander Control Data	140
	Program Guide	10
	Closed Captioning	10
	Program Mode Control	2
	Conditional Access	400
PROGRAM UNRELATED SERVICES	Teletext Services	
	Other Digital Services	
	Overhead	To be determined by System Proponent
Total Bits		1458

High Definition Television refers to an extremely sharp picture, with nearly twice as many lines as the current broadcast standard, National Television Systems Committee (NTSC) video; but ATV and HDTV will also be based on digital sound technology, which is what makes compact disc players sound so crisp and clear. Digital audio also makes it practical to compress sound data so that the bandwidth allocated to the sound portion of a TV channel can carry more sound tracks. When audio data compression becomes standard for television sound, it will make allocating bandwidth for video description easier. Until then, the SAP channel is accessible to tens of millions of homes now, giving it a big head start over competing technologies for receiving DVS.

The FCC is expected to set a standard for HDTV/ATV by September 30, 1993, but it is expected to be many years before there is a sizeable audience, given the cost of the initial sets is estimated to be $2000-$3000. It will be far longer before existing televisions become obsolete, because the ATV standard will be downward-compatible with existing televisions, so an interim solution for providing DV cannot be viewed as a solution to a problem that is about to solve itself. ATV is still over the horizon, and 200 million non-ATV television receivers will be around for a long time.

EIA Multiport Standard

The EIA Multiport Standard is intended to provide a standard for seamlessly connecting external equipment to a television, such as a closed captioning decoder, or cable TV without a separate cable box. It should be investigated as a possible standard for connecting any special devices required for receiving video description.

10.0 COST CONSIDERATIONS OF ADVANCED DESCRIPTIVE VIDEO TECHNOLOGY

Based on the Smith-Kettlewell and COSMOS reports, it appears that the technical and regulatory environments offer no serious obstacles to the provision of descriptive video (DV). However commercial viability, whether the cost of providing the service will be offset by a sufficiently large number of users to justify the cost without subsidies, is questionable, depending on marketing as much as cost.

According to the EIA, most households are now connected to a cable system. The EIA estimates that the major commercial networks, which once had as much as 90% of the prime time TV audience, now have about 67%, with 24% going to the cable networks. The big commercial networks, especially CBS, have been hit hard by that loss of market share, and it may be appropriate to get the cable companies involved in helping to defray the costs of described video, since they do not pay royalties to retransmit commercial stations' programs.

The costs of descriptive video services fall into two primary categories: those incurred by the provider of the services and those incurred by the user. The costs to the provider include network equipment modification or adaptation, adaptation of existing computer equipment to compose narration to fit the programs, labor costs for creating the narration, the costs of coordinating with production studios, and finally, the cost of upgrading local (affiliate) station equipment to enable them to broadcast DV.

This section concentrates on the cost of producing DV, and of broadcasting it to homes over the SAP channel. Other broadcasting options exist, but the SAP channel is presently the only practical medium available that does not separate audio tuning from video tuning. It is also the only transmission channel for which detailed cost studies are available. The reports reviewed are in essential agreement regarding these costs and they present a range from the cost of providing a low level of "special" programming through the cost of providing a level of service comparable to current closed captioning. Table 7 displays those costs. The "Low Cost" assumes 2 hours of programming/week; the "High Cost" assumes 50 hours of programming, a level similar to that of closed captioning programming. The cost of coordinating DV with the production studio is listed as "Hard to Quantify," but time is money; although the cost may not be easily quantifiable, it must be recognized.

Table 7. Costs of Described Video

Cost Factor	Low Cost	High Cost
Network Modification/Adaptation	$1,000,000 for up to 2 hours/week	$10,000,000 to $20,000,000 to upgrade network
Adapting Computer Equipment to Compose and Create Narration	$800 for labor-intensive computer program	$10,000 for full editing system
Labor Cost for Producing Narration	$4,500/week for 2 hours/week	$112,500/week for 50 hours/week
Coordination w/Production Studio	Hard to Quantify	Hard to Quantify
Affiliate Station Upgrade (Equipment)	Up to $20,000 but probably much less	Up to $20,000 but probably much less
Affiliate Station Upgrade (Labor)	Depends on Station Layout	Depends on Station Layout
Total Recurring Production Cost (Labor Plus Recording Studio Usage)	Approximately $3000/program hr	

The cost to upgrade an affiliate station's equipment will vary from station to station. The station's main transmitter may need a SAP generator for $4,500 to $8,000, but it may have come with one or have lower-cost modules available. Some stations also require modification of a satellite link for $250 to $1,050, an audio console for $2,000-10,000, modification of a microwave link for about $2,000, or even a new transmitter, though the old one would probably have been replaced within five years anyway. On top of that, there are labor costs, which depend upon station layout.

Obviously, the costs are not additive. A simplistic comparison of costs per hour of programming can be made using the cost numbers provided by the two studies. If one makes the assumption, under a low cost scenario, that over a three year period the network would provide two hours of programming each week, the cost would be $1,000,000 for the equipment adaptation, $800 for computer modifications, and $702,000 labor costs for creating the narrative for a total of $1,702,800 or approximately $5,500 per hour of programming. Under the high cost scenario, the network equipment modification would be a maximum of $20,000,000, with $10,000 for computer equipment and $17,550,000 labor costs to create the narrative for 50 hours of programming each week for three years. This is a total of $37,560,000 or approximately $4,800 per hour of programming. In each case, once the initial high cost of adapting or modifying the network equipment is absorbed, the network's cost of providing DV is limited to the labor cost involved in creating the narration and coordinating with production studios, the marginal cost of maintaining the equipment, plus the cost of studio time (non-labor).

Based on these reports, equipping an affiliate station with a satellite downlink upgrade, audio console, modified microwave link, and SAP generator would cost a maximum of $20,000 per affiliate, without labor costs. Under the low cost scenario of two hours of programming per week for three years, this would equate to a cost of $64 per hour of programming; under the second scenario, if the maximum of 50 hours of programming were broadcast over the affiliate station, the cost per hour of programming would be less than three dollars.

The cost to the consumer of receiving DV programming is limited to the cost of the receiver. According to each of the reports reviewed, the cost equates to the marginal cost of purchasing a television equipped to receive stereo broadcasts which is about $150. Other alternatives exist sporadically, including a $50 device to equip an existing TV with a SAP decoder that picks up the station to which the TV is tuned via a coil. It turns out that shielding to reduce interference to and from TVs and VCRs is actually a disadvantage for such a system. That system sometimes requires installing the coil inside the TV, but other retrofit systems, costing less than $150, act as the TV tuner. Unfortunately, acting as a tuner means bypassing the TVs remote control. The cost of a radio to receive the SAP channel would be expected to be on the order of $50, in consumer quantities. The issue of commercial viability hinges on whether there are a sufficient number of viewers, visually impaired or otherwise, who will both purchase the receiver and view the DVS programs. According to network representatives interviewed as part of the COSMOS study, approximately one million new viewers are needed to justify spending network funds on a new product or service. Numerically, this need should be met if ten percent of the visually impaired population become viewers or if a significant non-visually-impaired market is found. Whether this criterion can be met depends on the number of affiliate stations broadcasting the programming, the size of the population at which the programming is targeted, and the receptiveness of the population.

Experience with closed captioning suggests that the size of the audience may be significantly smaller than the ten percent required. The COSMOS study indicates that even though the hearing impaired population is larger than the visually impaired population, closed captioning viewership has never reached the one million needed for commercial viability. It is therefore possible that some form of subsidy will be necessary to introduce described video to commercial television, especially since it is more expensive to produce than closed captioning.

Two issues not sufficiently addressed in the COSMOS study of commercial viability are the differences between DV and closed captioning and the possibility of marketing DV to the non-visually-impaired population. Both of these considerations could make DV more commercially viable than the COSMOS study estimated.

Visual impairment is arguably less of a barrier to hearing about the benefits of DV than hearing impairment is to discovering the benefits of closed captioning. Also, receiving closed captioning has, until 1993 at least, required the purchase of a special decoder box. Anyone purchasing such a box must be willing to recognize and accept his hearing impairment and to take the trouble to find out where closed captioning decoders are available. Described video on the SAP channel can be received with equipment used by anyone and available in any department store. Visually impaired people are used to listening to the radio; severely hearing impaired people may not have been used to watching TV.

Marketing described video as a service to everyone, not only people with visual impairments, could also make it commercially viable. For example, the video description tracks for TV shows could also be used as radio shows, following the model of Cable News Network (CNN) broadcasts. Thus, DV could provide a source of revenue for the TV networks, easily meeting the 1,000,000-listener criterion in the Los Angeles, Washington, and New York markets alone. (WGBH may also want to distribute their video description as radio shows.) Everyone can benefit from described video because many TV shows, the classic example being detective shows, are visually hard to follow. Also, people often listen to programs while doing other things. Perhaps the issue is not whether there is enough of a market for DV but rather how to find it, although marketing is beyond the scope of this study. Marketing foresight may be the key to making DVS more commercially viable. Closed captioning was heavily subsidized because it did not have a large market. DVS should be marketed to a broader audience; that way it is associated with ability, not impairment. According to *USA Today* (July 11, 1990, p. 6D), "forty percent of the 60,000 closed-captioning decoders sold in 1989 were to people for whom English is a second language." That market was not anticipated when closed captioning was initiated, and it points out the importance of looking at alternative markets for products. Capitalism ensures better quality at lower per-user cost when a product is in higher demand. That generalization should apply to stereo TVs, VCRs, and SAP-capable TV radios, as well as DV production.

11.0 COST BENEFITS TO PERSONS WITH SENSORY IMPAIRMENTS WITH EARLY INCLUSION OF DESCRIPTIVE VIDEO

Persons with sensory impairments will benefit from early inclusion of descriptive video in several ways. These benefits essentially come out of a Government decision to support DV. At present, the second audio program (SAP) channel is the obvious choice for delivery to the consumer. The vertical blanking interval (VBI) appears to be the most promising option for getting DV to commercial networks' local affiliate stations. Consumers need not worry about how that is done.

Given a decision to produce described video on a broad scale over the SAP channel, through both PBS and commercial television networks, persons with visual impairments (and, not incidentally, everyone else) would then be able to buy products capable of receiving the SAP channel whenever they replace or upgrade their TVs, VCRs, and/or TV radios, confident that money would be well spent. That will reduce the need for retrofitting equipment, which always costs consumers money, performance and features.

At the same time, manufacturers need to add a wider selection of SAP-capable receivers to their product lines, including SAP-capable portable radios and probably car radios. The present selection of SAP-capable receivers available to consumers is completely inadequate for the more severely impaired end of the visually impaired population and marginal (generally too expensive) for the rest. If broad consumer demand could produce a sufficient market, the visually impaired population would benefit through lower product prices and better selection.

Finally, given the prospect of Government subsidies as needed, producers, networks and affiliate stations should become more willing to produce described programs, making consumers more willing to invest in equipment, thus stimulating the market.

Of course, it is important to watch out for the "chicken-and-egg" effect here. Producers, broadcasters and equipment manufacturers must take the lead in producing described video, arranging to distribute it, and developing a wider range of equipment to receive video description or consumers will not buy the receiving equipment. If everyone waits for someone else to take the lead, described video could fizzle or be very slow in coming.

12.0 PRESENT GOVERNMENT INVOLVEMENT IN DESCRIPTIVE VIDEO TECHNOLOGY

The Government is currently supporting described video in three ways:

- The Department of Education Office of Special Education Programs (OSEP) is supporting described video on PBS, primarily through the second audio program (SAP) channel, but also through Radio Reader Services.

- OSEP is also supporting described video on videotape.
- The Department of Education's NIDRR is sponsoring research on transmitting video description on the TV vertical blanking interval (VBI).

The FCC has no regulations that apply exclusively to DV, but FCC regulations permit stations to broadcast DV if they choose to do so. No protected standards mandate the use of any channel exclusively for DV.

13.0 DESCRIPTIVE VIDEO TECHNOLOGY TIMELINE

PBS is now broadcasting eight described series over the SAP channel on 61 TV stations, with 14 Radio Reading Services providing an alternative or a backup for areas that do not have SAP-capable stations. At least one small private cable network broadcasts "classic" movies with descriptions on its main audio channel. The major commercial television networks do not provide video description.

With the assistance of the Department of Education, WGBH has continued to expand its production of described video, adding more series. However, it appears that Government action, in the form of subsidies, will be needed to get commercial stations involved with producing and distributing described video.

At present, for transmitting DV from network affiliate stations to homes, the SAP channel is the only practical medium that does not separate audio tuning from video tuning. Unfortunately, the equipment and labor costs of adapting an entire network to transmit DV are presently high, due primarily to the cost and potential complexity of handling an extra audio channel at the network facility.

The most promising solution to the network facility problem is to distribute the extra audio channel over the VBI of the network's video signal. That way, the network facility remains intact, and the extra audio channel is inherently sent wherever the picture goes. Affiliate stations can then decode the VBI signal and route the resulting audio onto the SAP channel at the facility where they normally add subcarriers: either at the affiliate station or at that station's transmitter site. Development and testing will be required for such a system, so it will take two or three years before it becomes available. In the mean time, commercial stations should be able to make way for SAP capability by replacing transmitters that are not SAP-capable as they modernize their facilities.

Advanced television systems may be on the market as early as 1993 or 1994, and it is imperative that they incorporate an audio channel dedicated to DV. However, they may not begin to make existing televisions obsolete until after the turn of the century.

14.0 PROPOSED ROAD MAP FOR INCLUSION OF DESCRIPTIVE VIDEO CAPABILITIES

The COSMOS study of the commercial viability of DV concluded that "supportive" marketing conditions would be needed for DV to be produced and distributed by the commercial networks. That would include both startup and production investments, legislation, and FCC regulations. An important role of the Department of Education would be to ensure that alternatives to the SAP channel are considered, but enough direction must be provided to ensure that the market for DV receiving equipment will not be diluted by multiple incompatible technologies. In a small market, multiple solutions are generally no solution, the only exception being if they are either compatible or if one solution can be used by everyone and the other is provided as a higher-cost option. SAP would be a good common solution, but transmitting extra audio channels over the VBI would represent a good higher-cost option to the consumer who can afford to pay on the order of $200 for the receiver. Stations might preempt DV over the SAP channel with second-language broadcasts in some parts of the country, but SAP would be most likely to become a popular consumer commodity; special DV decoders (using VBI, for example) would add an expense and be less convenient to use, but they would provide an additional audio channel. As people move toward cable TV, it might make sense to offer cable boxes that handle the extra audio channels as a one-time-expense option. Of course, cable companies would presumably prefer to charge by the month for that service, but it may be necessary to regulate that since the cost to the cable company would really be a one-time cost. Given the growing role of the cable companies in delivering TV to homes, consultation with the cable networks, cable system operations and cable equipment manufacturers must be an integral part of efforts to bring DV to both commercial television and public television. The general rule for delivering DV should be to give all networks maximum flexibility in how they handle the audio for DV internally, but ensure that consumers need only invest in one type of receiver, preferably one with a broad consumer market (SAP).

Although ATV/HDTV will not make existing televisions obsolete in the next few years, prompt action is essential to ensure that the standards for ATV/HDTV, scheduled to be adopted in September 1993, incorporate a channel reserved for DV.

COSMOS also recommended conducting tests to find out who the DV audience will be. Although the COSMOS study did not consider the issue of the non-visually-impaired population using DV, that issue may be critical to the commercial viability of DV. It certainly affects how much DV needs to be subsidized. It is always easier to ensure the existence of a service if it is in higher demand. Careful marketing could be absolutely critical to the future success of DV. One option might be to test market DV over the National Public radio network, for the sighted audience. The results of such a market test could be used to predict the commercial viability of DV service.

For producers, a general guideline for developing DV technology is to ensure that nothing prevents later advances. Specifically, Smith-Kettlewell suggests keeping the following audio tracks on a four-track recording of the audio for described video: 1) the time code, 2) the voice of the narrator alone, 3) the original sound track, and 4) the mix of the narrator and sound track. They also recommend keeping a computer disk containing the narration text and corresponding time code in case the text is to be transmitted to a speech synthesizer instead of the voice of someone reading it. Finally, a paper or electronic copy of the script, with time code notations, should be archived. The progress of DV would be most rapid if the format of this information is standardized as soon as a consensus can be reached.

15.0 POTENTIAL PROGRAM SCHEDULE

Figure 2 presents a potential program schedule for the development of descriptive video using advanced technologies. The Department of Education could act as the program administrator providing basic research and development funding. Within three to five years, alternative transmission techniques could be provided for descriptive video transmission.

	92	93	94	95	96
VBI Research	X	X	X		
Fund Public Broadcasting	X	X	X	X	X
Support ATV/HDTV DV Stds	X	X	X		
Establish a Focus Group on DV Market Appeal	X				
Perform DV Market Studies for General Market		X			
Fund Research on DV Services to Visually Impaired	X	X	X		
Begin Subsidize Network DV Using VBI			X	X	X
Cable TV Investigations	X	X	X		
Fund ATV/HDTV DV Research			X	X	X
Subsidize ATV/HDTV Broadcasts					X
Training/Education/Publicity for DV			X		

Figure 2. Proposed Schedule

BIBLIOGRAPHY

Unpublished text, phone conversations and visit with:
Descriptive Video Service
WGBH
125 Western Avenue
Boston, MA 02134
(617) 492-2777 ext. 3490

Technical Viability of Descriptive Video Services
William A. Gerrey
Rehabilitation Engineering Center
The Smith-Kettlewell Eye Research Institute
June 1990
U.S. Department of Education
Award No. H133E00004-90A
(415) 561-1677

Commercial Viability of Descriptive Video Services
K. Zantal-Wiener, R.K. Yin, R.T. Yin, S.E. Wiley
COSMOS Corporation
May 1990
U.S. Dept. of Education, Office of Special Education Programs
Contract No. HS89021001
(202) 728-3939

Phone Conversation and Brochures by Eric Small
Modulation Sciences, Inc.
12A World's Fair Drive
Somerset, NJ 08873
(908) 302-3090
(800) 826-2603
10/16/91

Office of Science and Technology
OST Bulletin No. 60
Multichannel Television Sound Transmission and Audio Processing Requirements for the BTSC System AND Amendments to Rules and Regulations for Multichannel Television Sound

Television Engineering Handbook
K. Blair Benson
Engineering Consultant, Editor in Chief
Mc Graw-Hill Book Company
New York, NY
1986

Phone Conversation with Mr. Kelly Williams
Television Engineer
National Association of Broadcasters (NAB)
(202) 429-5300

Proposal for TV Audio & Data Services for Simulcast ATV Systems
May 13, 1991
From WGBH but based on "original version submitted to the S3 Committee by T.B. Keller,
Consultant CableLabs"

Metropolitan Washington Ear
Ms. Gwen Garfinkel
(301) 681-6636

V8600/1 Speech Synthesizers Data Book
RC Systems
121 West Wine Sap Road
Bothell, WA 98012
(206) 672-6909

Reference Data for Engineers: Radio, Electronics, Computer, and Communications
Seventh Edition
Howard W. Sams & Co., Inc.
Indianapolis
1985

Ferrel G. Stremler
Introduction to Communication Systems
Second Edition
Addison-Wesley Publishing Company
Menlo Park, CA
1982

Television Vertical Interval Encoding and Decoding
EEG Enterprises, Inc.
1 Rome Street
Farmingdale, NY 11735
(516) 293-7472

Consumer Electronics Annual Review
Electronic Industries Association
Consumer Electronics Group
Washington, DC
1990

Consumer Electronics U.S. Sales
Electronic Industries Association
Consumer Electronics Group
Washington, DC
June 1990

Britannica Book of the Year, 1991
Encyclopaedia Britannica, Inc.
Chicago
1991
p. 360

That All May Read
National Library Service for the Blind and Physically Handicapped
The Library of Congress
Washington, DC
1983

ADAPTIVE MODEMS AND

TELECOMMUNICATIONS DEVICES FOR THE DEAF (TDD)

MARCH 1992

Prepared by

Daniel E. Hinton, Sr., Principal Investigator
and
Charles Connolly

SCIENCE APPLICATIONS INTERNATIONAL CORPORATION
3701 N. Fairfax Drive, Suite 1001
Arlington, VA 22203
(703) 351-7755

1.0 SCENARIO

Adaptive Modems and Telecommunications Devices for the Deaf (TDD)

2.0 CATEGORY OF IMPAIRMENTS

Persons with Hearing Impairments.

3.0 TARGET AUDIENCE

Consumers with Hearing Impairments. Persons with sensory impairments will benefit from enhanced access to media information and telecommunications services. This scenario on advanced modem technology will provide a means to disseminate information to consumers with hearing impairments. It will provide consumers with hearing impairments a better understanding of modem capabilities beyond the slow--45.5 bits per second (bps)--Baudot TDD modems they are now using. In addition, this scenario transition from existing Baudot TDD modems to full ASCII modems operating at 1,200 to 9,600 bps proposes methodology.

Policy makers, including national representatives, Government department heads, and special interest organizations. Policy makers can use this scenario to better understand the issues related to telecommunications access for persons with hearing impairments. They may also use it as a point of departure to understand how advanced modem technology is making it possible to use legislation or regulatory action to mandate the inclusion of Baudot TDD access requirements in telecommunications modems.

Researchers and Developers. The R&D community will benefit from this scenario through a better understanding of the Baudot TDD communication needs of persons with sensory impairments. Better understanding of TDD requirements assists researchers and developers in designing TDD functions into future products and promoting an environment in which the needs of persons with hearing impairments are met.

Manufacturers. Manufacturers will benefit through a better understanding of the potential market size and the existing Federal Government requirements for telecommunications access which can be met by adding TDD capability to their modems.

4.0 THE TECHNOLOGY

New advanced microchip modem technology offers a leap forward in design flexibility over existing modems. Advanced modem technology now makes it possible to implement TDD modem functions in advanced ASCII modems. This may be accomplished through software resident on the modem chips or on the host computer system. Either way, expensive hardware modifications are not needed because the advanced modem technology uses digital signal processors, programmed for the modem tone generation and detection functions previously accomplished using expensive hardware.

An example of this technology is the Rockwell RC9696/12 modem chip set which uses a digital signal processor that can be programmed for dual tone operation from a host computer. Therefore, the Baudot TDD tone detection and generation could be included in or added to the modem via software rather than a hardware redesign even for the 9,600 bps modems presently sold that use this modem chip set.

5.0 STATEMENT OF THE PROBLEM

Advanced telecommunications modems have been developed to meet the American consumer's needs for high quality data transmission at 9,600 bps over standard telephone lines. At this time, access to this new technology by persons with sensory impairments is not being addressed by Government or industry (i.e., management, researchers, or marketers). This could perpetuate a situation in which persons with sensory impairments who use Baudot TDD modems have little or no access to telecommunications services (i.e., electronic mail, database retrieval systems, and person-to-person communications). Unless action is taken, this barrier could persist into the foreseeable future.

6.0 DEPARTMENT OF EDUCATION'S PRESENT COMMITMENT AND INVESTMENT

The Department of Education has funded TDD modem research and development over the past 20 years. With the advent of personal computers in 1975, they began to fund research and development of dual capable Baudot TDD and ASCII computer modems specifically targeted for persons with hearing impairments. Presently, the development of Baudot TDD and TDD-compatible ASCII modems is a stated research priority of the Department of Education as follows:

- In the Department of Education's "Small Business Innovative Research (SBIR) Program Phase I Request for Proposal," issued January 11, 1991, research topics related to TDD modem access included:

    1. Adaptation or development of commercial quality integrated voice/ASCII/Baudot teletype-payphone units.

    2. Adaptation or development of an inexpensive modem add-on device to enable ASCII modems to communicate with Baudot TDDs.

    3. Adaptation or development of an add-on controller which will enable telephone switching devices to automatically recognize incoming Baudot TDD calls and switch them to the correct device--a capability

which is currently available for FAX, ASCII, and voice calls that come in on the same telephone line.

- The findings of the National Workshop on Rehabilitation Technology, a cooperative effort of the Electronic Industries Foundation (EIF) and the Department of Education's National Institute on Disability and Rehabilitation Research (NIDRR), indicated a need for research to develop computerized techniques to facilitate the use of telephone systems and broadcast media by deaf, hard-of-hearing, and visually impaired/hearing impaired persons, including voice/Baudot TDD interfaces. The workshop's Consensus Panel recommendations included modem standards: "Concerns are incompatibility of Baudot TDDs with standard computer modems used for information transfer, the need for interfacing existing Baudot TDD units with modern computers, and the lack of standards, specifications, and protocols for TDD-compatible standard modems.

- The *Federal Register*, December 4, 1990, states the Final Funding Priorities for the NIDRR for fiscal years 1991-1992. These priorities include "creating more accessible communication environments for (the deaf and hard of hearing) population." One of the stated approaches to meeting that goal is to "conduct at least one national study of the state of the art to identify current knowledge and recommend future research."

The program, titled "Examining Advanced Technologies for Benefits to Persons with Sensory Impairments," conducted by the U.S. Department of Education who developed this scenario, represents one such study. The Panel of Experts for this program included nationally known experts in technology and persons with sensory impairments. When the Panel met on February 7-8, 1991, there was a consensus that adaptive modems and TDDs are the number one advanced technology priority for people with hearing impairments for three reasons. First, a Baudot TDD capability for advanced modems would substantially impact telephone and telecommunications media access for the hearing impaired. Second, there are no technological obstacles to making Baudot TDD capability standard on all new advanced modems with minimum development time. Finally, the number of advanced technology modems in use is relatively small, but growing fast, as Figure 1 shows. An additional fall out from meeting the needs of persons with hearing impairments is that for the first time the public will have direct telephone access to persons with hearing impairments on a large scale.

TDD modem access is a priority because there are an estimated 400,000 Baudot TDDs in use in the United States; a country with over 30 million hearing impaired people. However, the computer modems being used on bulletin boards, Government retrieval systems, 911 emergency services, and home shopping networks, to name only a few, employ ASCII modems that cannot be used with Baudot TDDs. Access to computer modem

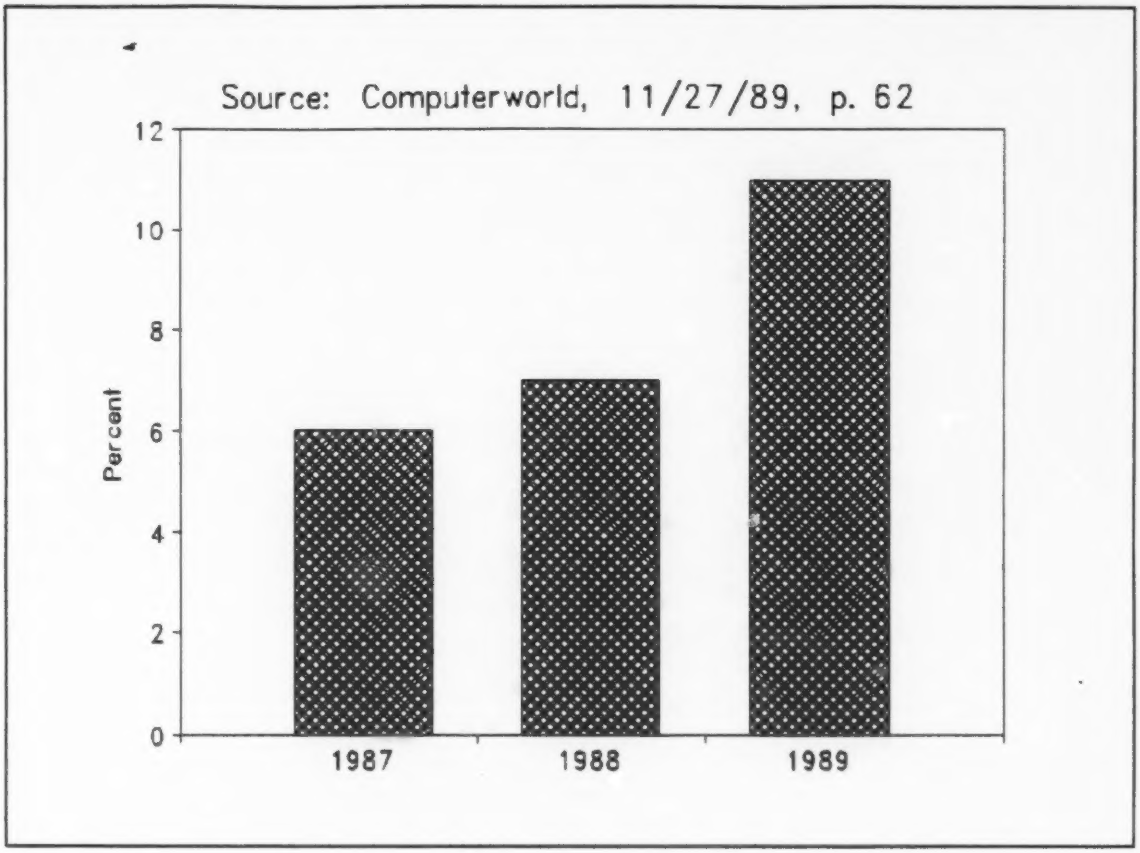

Source: Computerworld, 11/27/89, p. 62

Figure 1. Dial-Up Modems at 9600 BPS or Higher

technology via Baudot TDDs has been limited to specially designed modems, due to the implementation of modem tone detection and generation functions in hardware. The Department of Education funded several of these TDD/ASCII modems that have a maximum data rate of 1,200 to 2,400 bps in the ASCII mode and 45.5 bps in the Baudot TDD mode. Table 1 is a representative list of existing TDDs. As advanced modem technology is applied, it will be necessary to either develop new limited market modems to meet the ever changing market, incorporate TDD modem functions into all advanced modem technology, or develop standards that require adding ASCII capability to all new TDDs.

The main reason for not including a Baudot TDD function in ASCII modems has been that the manufacturers developed unique hardware and software for each new generation of modem for the general computer market. Backward compatibility with earlier ASCII modems was accomplished using hardware and software capable of receiving or

Table 1. Telephone Devices for the Deaf (TDDs)

TDD	Company	ASCII	ASCII Baud Rate	Cost Without ASCII	Cost with ASCII
Touchtalk Travelpro	ZiCom Tech.	opt.	300	$179	$229
Superprint ES	Ultratec	yes	300		$635
300 Zi Modem Card	ZiCom Tech.	yes	110, 300 1200,2400		$289
1310+ Terminal	AT&T	yes	110, 300		$330-485
Pay Phone TDD	Ultratec	opt.	300	$80/mo.	$80/mo.
Supercom	Ultratec	opt.	300	$299	$349
SSI 240	SSI	no		$590	
PV20 Series	Krown Res.	opt.	300	$219-249	$305
CM-4 Modem	Phone TTY	yes	110, 300		$349
Minicom	Ultratec	no		$199-229	
TE 98	Auditory Display	no		$275	
SM85	Krown Res.	yes	300		$349
LUV1	Amer. Comm.	no		$149	
InteleModem	Weitbrecht	yes	110, 300		$289
MP20	Krown Res.	no		$449-499	
MP20D	Krown Res.	yes	300		$579
Superprint 100,200,400	Ultratec	opt.	300	$359-499	$409-549
Compact	Ultratec	opt.	110, 300	$289	$349
PCT	Trident	yes	1200		$595

transmitting only a narrow range of frequencies. Baudot TDD-compatible modems were not implemented because the manufacturers would have had to add expensive hardware--in the form of electronic chips--to the system to generate and detect the unique Baudot TDD tones. The market for Baudot TDD modem capability was considered to be insignificant when compared to the number of general computer modems installed in the U.S.; shown in Figure 2. In addition, Baudot TDD modems do not use a carrier detection scheme so it is not possible for the hardware to automatically determine, with certainty, if a Baudot TDD or voice user is present.

Sources: Statistical Abstract of the United States, 1990, p. 759; PC Week, April 15, 1991, p. 113.
The curves were extrapolated from data points from these sources, marked with an X.

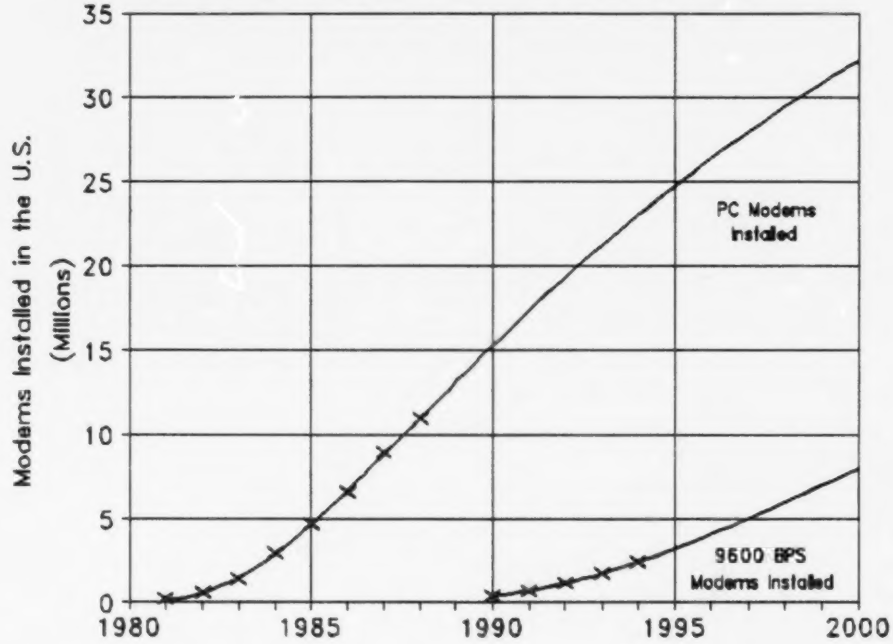


Figure 2. Modems Installed in the U.S.

7.0 ACCESS TO TELECOMMUNICATIONS INFORMATION AND COMMUNICATION MEDIA

Many federal, state, and local laws influence telecommunications for hearing impaired people, just as these laws influence telecommunications for the general population. The most important single law related to telecommunications for hearing impaired people is Public Law 101-336, enacted July 26, 1990. Better known as the Americans with Disabilities Act (ADA), this law has broad implications for all disabled Americans.

Title IV of ADA relates to telecommunications relay services for hearing impaired and speech impaired individuals. It modifies Title II of the Telecommunications Act of 1934 (47 U.S.C. 201 et seq.) by adding Section 225. This section provides that each common carrier providing voice transmission services must also provide telecommunications relay services for hearing-impaired and speech-impaired individuals within three years of the enactment of ADA.

Within one year of the enactment of ADA (July 26, 1991), the Federal Communications Commission (FCC) must prescribe regulations which:

a) Establish functional requirements and guidelines.
b) Establish minimum standards for service.
c) Require 24 hour per day operation.
d) Require users to pay no more than equivalent voice services.
e) Prohibit refusing calls or limiting length of calls.
f) Prohibit disclosure of relay call content or keeping records of content.
g) Prohibit operators from altering a relayed conversation.

The FCC must ensure that the regulations encourage the use of existing technology and do not discourage or impair the development of improved technology.

The national relay service will probably involve several hundred million calls a year and will be extremely expensive. Any development which shaves a few seconds off an operator's time on a call will mean significant long term monetary savings. This puts tremendous pressure on the telephone industry to develop an efficient technologically advanced service. Since ASCII transmission is so much faster than Baudot transmission, there should be a strong move toward converting from Baudot to ASCII technology. ADA encourages "the use of existing technology" and the current Baudot TDD system is mostly Baudot-based; thus a bridge between Baudot and ASCII equipment is required.

8.0 POTENTIAL ACCESS IMPROVEMENTS WITH ADVANCED MODEM TECHNOLOGY

Advanced technology modem chip sets implement all modem functions in software on the chip sets, so all the modem manufacturers will have to do is write the user interface software or proprietary system functions. This includes the screen display formats, routines to save files that are received, and routines to send information files to the modem. The advanced modem chips have the capability to distinguish between data and voice, although they do not yet have word recognition capability. The advanced modem chip sets can also be programmed to emulate any dual-tone modem, such as the Baudot TDD modem function. The simple emulation program may be included on the chip set, resident on the host computer, or downloaded from the host computer into the modem chips memory.

Table 2 is an extract from the AT&T's WE DSP16A-V32 modem chip set specification and is representative of capabilities offered by other advanced modem chip set. This is typical of what will be incorporated into modems over the next one to five years. Rockwell's popular 9,600 bps RC9696/12 chip set is similar. It lacks a sleep mode, requires more power (1.9W typical, 3.5W maximum), lacks 4-wire operation and a secondary channel. It has much less built-in echo cancellation capability (53.3 msec), and is mainly set up for connecting to 8086-based computers and the CCITT V.24 interface. However, it was one of the first on the market and a number of manufacturers are presently using this modem chip set.

This advanced modem technology offers the potential for dramatic improvements in telecommunications access for persons with sensory impairments, using their existing Baudot TDD modems to access:

- Databases
- electronic mail systems
- bulletin board systems
- mail order systems
- other modem users.

It is critical to recognize, however, that these improvements only come if the new modems support Baudot TDD access. Until then, although advanced modems may be easier to retrofit for Baudot TDD, the vast majority of modems will still be a barrier to improved media access for the hearing impaired who do not have ASCII-capable modems.

The key is that these services will be among the first to use this advanced technology to serve a broad segment of the general population. These modems are backward compatible with most of the other modems because the advanced modem chip sets are able to distinguish the various modem formats and automatically configure themselves for the appropriate mode of transmission (Table 2). With a Baudot TDD mode added as part of an enhanced instruction set, or as an externally programmable feature, any person with a Baudot TDD modem could access the systems discussed above, given software was added to allow the information to be displayed in a Baudot TDD compatible mode.

Table 2. Advanced Technology Modem Features

- Compatibilities:
    - CCITT V.32: 9600 (TCM), 9600 (QAM), 4800 (QAM)
    - CCITT V.22bis: 2400 (QAM), 1200 (QAM)
    - CCITT V.22: 1200 (QAM), 600 (QAM)
    - CCITT V.23: 1200 (FSK), 600 (FSK)
    - CCITT V.21: 300 (FSK)
    - Bell 212A: 1200 (QAM)
    - Bell 103: 300 (FSK)
- - 9600, 4800, 2400, 1200, 600, or 300 bps transmission speeds, plus a programmable speed of 0-300 bps
- Low power consumption:
    - Sleep mode power consumption under 100 mW
    - Typical active power consumption under 400 mW
    - Maximum active power consumption under 1 W
- Automode start-up
- 2-wire and 4-wire full-duplex operation
- Full-duplex asynchronous inband secondary channel (typically 150 bps)
- Auto-dial and auto-answer capability
- Echo cancellation:
    - Frequency-offset compensation in far-end canceler
    - 1.2 seconds of internal bulk delay RAM (no external memory required)
- Interfaces:
    - Configurable 8-bit microprocessor interface allows easy, glueless communication with multiplexed and non-multiplexed microprocessors
    - Full V.24 interface
    - Constellation interface
- Can be configured to be hardware compatible with a socket designed to support a Rockwell International R9696DP modem module
- Operating status and line quality information including receive signal parameter reporting
- V.13 simulated carrier control
- V.54 remote loop test
- Flexible I/O for V.42 support

An alternative approach to providing ASCII modem access for persons with hearing impairments would be to require all Baudot TDD devices built after a specified date to be ASCII modem-compatible at the user level. This approach would specify a time frame in which all Baudot TDD modems for the hearing impaired would be converted to ASCII modem capability.

Several problems are encountered with this approach:

- A simple telephone handset connection to the modem would have to be replaced with a telephone RJ-11 jack connection to allow modem implementations of 1,200 bps and faster.
- The modem increases in price because the hardware for the Baudot TDD modem becomes more complex and expensive. For example, a microprocessor would be required in an ASCII-compatible TDD.
- The user must select the mode of operation prior to connection, or the machine must be smart. The advanced technology modem chip sets are smart enough to perform this function.

The advantage to making all new Baudot TDDs ASCII-compatible is that persons with hearing impairments move up into the computer-compatible modem world with little or no impact on existing computer modems.

9.0 ADVANCED MODEM TECHNOLOGIES

Figure 3 shows the trend in modem technology. The earliest modems ran at 45.5 bps, which eventually became the TDD standard, but they advanced to 110, 300, 600, 1,200, 2,400, 4800, and finally 9,600 bps, while Baudot TDD's remained at 45.5 bps. Worse, Baudot TDD's still use Baudot code, an obsolete character representation that has long since been replaced by ASCII. The 45.5 bps modem speed was adequate for most person-to-person calls since a 60 word per minute typeset only achieves an equivalent data rate for a Baudot transmission of 45 bps. However, for other services such as file transfer, this rate must be increased to reduce time and cost.

The first modems implemented all functions in hardware, severely limiting their flexibility to deal with different standards. Generally, as digital data processing speeds increased, it became possible to perform some functions in software, increasing flexibility. First, the simplest controls were implemented in software, then protocols to allow modems to "handshake." Eventually, digital signal processing and fast microcontrollers implemented all functions to be performed in software, and advanced modems became a reality.

The number of companies that are developing--or have developed--advanced modem chip sets is growing rapidly. Currently, the companies that make 9,600 bps modem chip sets

Table 3. Examples of Advanced Modems Now Available

Modem	Price	Chip Set
ATI 9600etc/e	$499	Rockwell
Best Data Product's Smart One 9642X	$599	Rockwell
Computer Peripherals ViVa 9642e	$649	Rockwell
Digicom 9624LE+	$995	Digital Signal Proc.
Everex Evercom 96E+	$699	Rockwell
GVC SM-96	$599	Rockwell
Hayes Ultra 96	$1199	Phylon Communications
Intel 9600EX	$799	Intel
Microcom QX/4232hs	$899	Rockwell
Multi-Tech MultiModem V32	$1149	Rockwell
Practical Peripherals PM9600SA	$699	Rockwell
Telebit T1500	$1095	Rockwell
UDS FasTalk V.32/42b	$795	Proprietary
U.S. Robotics Courier V.32	$995	Rockwell
Ven-Tel 9600 Plus	$899	Rockwell
Zoom Telephonics Zoom/Modem V.32 Turbo	$599	Rockwell

factor of three to four times, making these modems more than 1,000 times as fast as a Baudot TDD in some applications.

As an example of an advanced modem chip set, the Rockwell RC9696/12 contains a modem data pump, a microcontroller, and microcontroller unit (MCU) firmware. The data pump consists of a digital signal processor--for all of the modem's signal processing functions--and analog circuitry. The microcontroller performs error correction, data compression, and protocols. Programs for the microcontroller are stored in a read-only memory (ROM), referred to as the MCU firmware. Although chip counts vary between manufacturers and are decreasing with time, the trend is toward this type of modem architecture. AT&T's chip set, for example, uses the same basic building blocks grouped into chips sets differently.

10.0 COST CONSIDERATIONS OF ADVANCED MODEM TECHNOLOGY

Figure 4 shows that advanced modem prices have been falling sharply in recent years. The TDD prices that were listed in Table 1 have been much more stable. It is projected that current top-of-the-line modems will be less expensive than TDD's within three years.

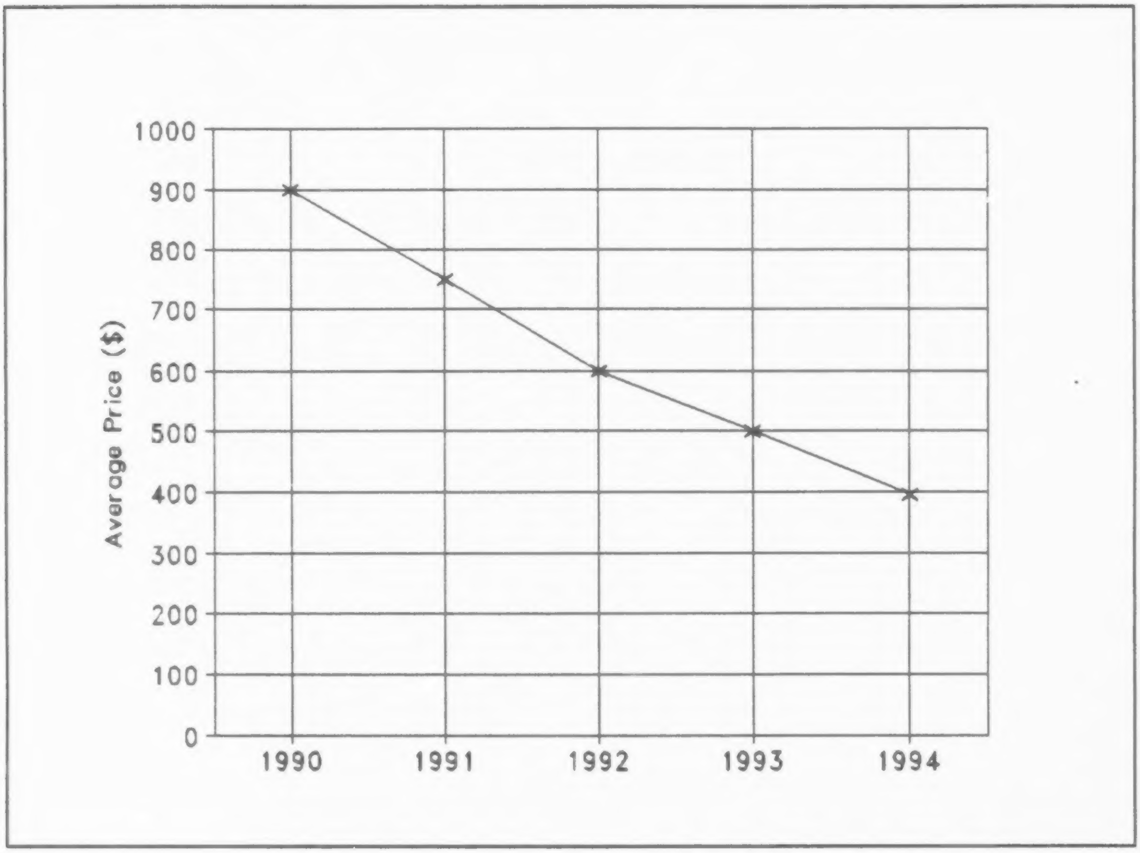


Figure 4. Advanced Modem Prices

Adding Baudot TDD function to advanced modems is estimated to cost $20,000 for each modem product line (the cost of adding about 100 lines of code to a program). If 100,000 units are sold, the total additional production per unit cost is about 20 cents. In practice, the cost of developing an inexpensive feature--like Baudot TDD capability--is generally absorbed in a short time. Novel features justify higher prices until the feature becomes standard and manufacturers have covered their costs. Maintaining the additional program lines to support Baudot TDD capability over the life cycle of a modem would add about a penny to the retail price of each modem. In short, the per unit cost associated with adding Baudot TDD capability to advanced technology digital signal processor based modems is small, but that cost provides broad access to hundreds of thousands of Americans.

Looking at the long term picture (5-15 years), this small cost also enables the deaf community to slowly transition from the outdated Baudot 45.5 baud standard and transition to the technology being employed within the consumer electronics market in about five years (Figure 3). Until recently it was hard to justify paying for a more expensive ASCII modem and computer when a TDD could do an acceptable job. Now, some Baudot TDDs actually cost more than state-of-the-art ASCII modems, and modems that run 50 to 100 times as fast as a Baudot TDD sells for a third of the base price of a Baudot TDD. In two or three years it is expected that state-of-the-art modems will cost less than Baudot TDD modems because the consumer demand and technology will continue to drive the price of ASCII modems down and leave TDD modems at their current price.

11.0 COST BENEFITS TO PERSONS WITH SENSORY IMPAIRMENTS OF EARLY INCLUSION OF BAUDOT TDD FUNCTIONS IN ADVANCED MODEMS

The retail per unit cost of adding Baudot TDD modem function to existing advanced technology chip sets or manufacturers' software is less than $1.00 per unit, based on a 5 to 1 cost multiplier. Each advanced technology modem would then be able to connect with the full range of ASCII and Baudot TDD modems. Persons with sensory impairments could continue to use their existing Baudot TDD modems until they need replacement. When purchasing a new modem, they would upgrade to an advanced technology modem that was compatible with both Baudot TDDs and ASCII modems.

Within five years most interactive computer services will use the advanced modems, including Government, industry and educational institutions. In addition, several million individuals will be using these modems nationwide. Figures 1 and 2 showed the rapid market share increase for advanced modems. The earlier a Baudot TDD standard is developed and required in all advanced technology modems, the less costly it will be to persons with sensory impairments because it makes Baudot TDD-capable modems a mainstream consumer product. The installed base of advanced modems will then ensure access via software. This upgrade promotes the Department of Education goals through the implementation of Baudot TDD capability in all modems.

12.0 PRESENT GOVERNMENT INVOLVEMENT IN ADVANCED MODEM TECHNOLOGY

The Department of Education funded two modem projects that are listed in the FY89 NIDRR Program Directory. Many early TDD modem developments were funded by NIDRR and Office of Special Education Programs (OSEPS). One project was entitled

"Integrated, Intelligent, and Interactive Telecommunication System for Hearing/Speech Impaired Persons." This Phase II project was awarded to Integrated Microcomputer Systems, Inc., Rockville, Maryland, and featured TTY/TDD and ASCII compatibility, "remote signal control, direct connection to the telephone system, and text-to-speech voice announcer."

A Field-Initiated Research project, entitled "Deaf-Blind Computer Terminal Interface," was awarded to SAIC in Arlington, Virginia, for the development of an acoustical modem interface between the Telebrailler, Microbrailler, TDD, IBM-PC compatibles, and the Commodore 64C.

13.0 ADVANCED MODEM TECHNOLOGY TIMELINE

In the mid- to late-1990's, the use of most computers will be user friendly and require no more computer literacy than today's Baudot TDDs, and, a TDD will cost more than a much more flexible high volume computer. This is already happening. By the year 2000, manufacturers will not be able to recover their costs when they try to sell new TDDs at prices competitive with computer equipment. Several hundred thousand people will still have traditional Baudot TDDs. The Department of Education should prepare for the transition by sponsoring (1) mandating Baudot capability on all new computer modems sold after 1995 or (2) mandating ASCII capability on all TDDs. The following paragraphs discuss the technology that can make either option possible.

The recent introduction of digital signal processors (DSPs) has made it feasible to include Baudot TDD capability in all modems. With DSPs, the inclusion of Baudot capability is possible through relatively inexpensive software changes inside the modem, instead of expensive added hardware. If all ASCII digital signal processor based modems are required to be Baudot TDD-capable, hearing impaired consumers could transition from the outdated Baudot TDD standard without abandoning anyone with Baudot TDDs. By 1995, advanced modems will be included with most new computers. If Baudot TDD compatibility is mandated for digital signal processor based modems, any person or company buying a computer would then have access to all Baudot TDDs. However, that capability would include high-speed access to remote computers, databases, electronic mail, electronic bulletin boards, shopping networks, and other modem users.

A TDD-capable advanced modem receiving a call from a TDD user should be designed to automatically configure itself for the TDD without the need for human intervention on part of its call set up routine. Through attrition, TDD-capable advanced modems would replace many TDD modems by the year 2000. TDDs and standard modems would then have the same access capabilities, within their respective speed and display limitations. These limitations should not be minimized however. They render a TDD inherently useless for applications that require substantial data transfer or a multi-line screen.

Nevertheless, universal TDD access would be of great benefit to TDD users. Requiring all modems sold starting in 1995 to have Baudot TDD capability would make TDD access almost universal around the turn of the century, at an estimated cost of under $1.00 per modem at point of sale. That cost would be borne by everyone purchasing a modem, rather than concentrating the cost on the hearing impaired population. By the turn of the century, traditional TDDs should be phased out and replaced by personal computers with the new TDD-compatible modems. At the same time, making TDD-compatible modems more affordable would influence many others to purchase the computer and TDD-compatible modem.

Requiring all new TDDs sold to be ASCII-capable could perpetuate the five year lag in hearing-impaired ASCII compatible modem technology because low product demand generally leads to slower development of a technology. It could also raise the minimum TDD price. On the other hand, if TDD manufacturers take advantage of the similarities between ASCII-capable TDDs and computers, prices could go down because of a more focused demand and the five year lag time could be reduced or eliminated. However, "plain" ASCII Baudot TDDs are becoming a poorer investment every year due to the advances in computer technology. ASCII-capable TDDs already exist, so, from a technical standpoint, ASCII capability could be required on TDDs in three to five years. This approach provides strong incentives for people with TDDs to transition to ASCII-capable modems. Many hearing impaired people are older Americans, and the sensory-impaired population has historically been underemployed. Making all new TDDs ASCII-capable could postpone the transition from ASCII-capable TDDs to personal computers. But it can be argued that fast ASCII-capable TDDs, designed to take advantage of their similarities to computers, may be simpler and more functional for persons with hearing impairments as a whole than computers, at least for older Americans who have neither the desire nor the skills to operate computers or who may even be afraid of computers.

TDDs without ASCII capability are locked into technology that became obsolete in the late 1960's. Meanwhile, the data transfer rate of advanced modems is doubling every two to three years. Figure 5 shows this trend. Based on historical modem speed trends, in 20 years, transfer rates doubled approximately every year. Modem speeds are rapidly approaching the theoretical limit of 25,000 bps for conventional telephone lines. It is doubtful that 25,000 bps will be achieved on an arbitrarily chosen telephone line, although leased lines already permit transmission rates near the theoretical limit of 25,000 bps. Recently, data compression technology has enabled data transfer at effective rates for ASCII data as high as 38,400 bps, using mass-produced advanced modems with 9,600 bps speeds and 4 to 1 compression ratios. Redundant elements, which are simply patterns within the data and exploited via compression algorithms, increase the effective data transfer rate by 4 times. Eventually, using data compression, effective transfer rates could go as high as 100,000.

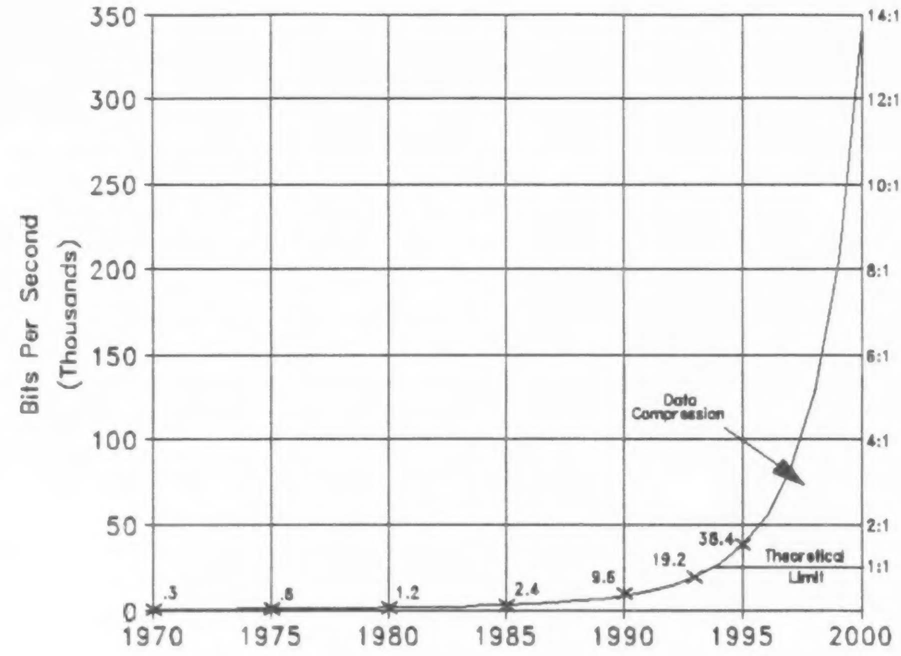


Figure 5. Projected Modem Transmission Rates

Currently, the number of advanced modems installed in the U.S. is doubling annually. Figure 2 showed this trend. Sales of newly-introduced products, like advanced modems, tend to level off with time as there are fewer new customers to be found, and the established base is being upgraded or replaced. For that reason, sales of 9,600 bps advanced technology modems are likely to level off around 1996. Also, 19,200 bps and faster modems will divert sales from 9,600 bps modems. Thus, referring back to Figure 2, the installed base of 9600 bps modems may be overestimated after 1994, although sales of faster digital signal processor based modems are expected to make up the difference. The long-range projections of the installed base of PC modems in Figure 2 should also be regarded as only rough estimates after 1988.

Regardless of the long-range projections given, the increasing doubling time of the installed base of advanced technology modems, and their increasing share in the modem market, represent a window of opportunity to transition from TDDs to ASCII based modem devices. As the base of advanced modems builds, the ability to make that generation of modems TDD-compatible--without large-scale retrofitting, which is unlikely to occur--decreases. Already, development time will prevent TDD-compatible modems from catching up to the state of the art until around 1995. Action by the Department of Education is required to ensure that the most heavily used modems are TDD-capable by the end of the 1990's. It is not cost effective to make older ASCII modems compatible with Baudot modems due to the necessity of modifying the hardware. Gradually these modems will be replaced with advanced technology digital signal processor modems. This is because faster modems mean more satisfied customers and employees, and lower phone bills, so businesses, Government offices, non-profit organizations and individuals will replace the modems that are most important for media access. That is the motivation for making the replacements TDD-capable as soon as possible. One major benefit of this transition would be a modem that could be used by emergency response units such as 911 services. Today, most 911 services do not have TDD or ASCII modem capabilities.

14.0 PROPOSED ROAD MAP FOR INCLUSION OF BAUDOT TDD CAPABILITY IN ADVANCED MODEM TECHNOLOGY

Competition in the modem industry will ensure that modem transmission rates double every two to three years, up to the theoretical limits of telephone lines and that data compression will push modem's efficient transfer rates well beyond 25,000 bps. Modem manufacturers have the economic incentive to use coding techniques, data compression, and other technologies to push beyond the theoretical limits. An initiative by the Department of Education would ensure that these modems will support ASCII and Baudot TDD operations. With research and development action over the next 3-5 years, TDD-capable modems could achieve parity with modems designed for the general public by the mid-1990's. If all new digital signal processing modems sold in the U.S. after 1995 are required to be TDD-capable, businesses and consumers would, for the first time, be able to buy a single-unit ASCII/Baudot modem at competitive prices, with the features and transmission rates available to all modem users. The cost of Baudot and ASCII TDD modems would be amortized over many thousands of units with the benefits of improved communication access for the hearing impaired and greater access to the hearing impaired by the general population. Three to five years after regulatory or legislation is enacted requiring all advanced technology modems to be TDD-capable, several million modems would have been replaced with Baudot TDD-capable modems and TDD access would be almost universal.

Table 4 shows an Advanced Modem Development Road Map. Before TDD compatibility of ASCII modems can be mandated, a standard for TDD modems must be established and technical requirements must be defined. The Department of Education has initiated a TDD modem specification committee through the Lexington Rehabilitation Engineering Center. This group includes representatives from Gallaudet, the Telecommunications for the Deaf, Inc., telephone industry associations and manufacturers. The Department of Education's next step should be to obtain a draft standard recommendation on TDD services to the telecommunications industry, the National Institute of Standards and Technology (NIST) and the FCC.

Table 4. **Advanced Modem Development Road Map**

Obtain Consultant Services
Form Department of Education Advanced Modem Development Committee
Define Baudot/ASCII Modem Requirements
Translate the Requirements into Technical Specifications
Generate Formal Standards
Conduct Engineering Studies
Begin Developing Baudot/ASCII Modems at 9600 BPS
Begin Developing Baudot/ASCII Modems at 19,200 BPS
Establish Points of Contact Between the Telecommunications Industry and the Department of Education to Ensure Future Dialog About Access Issues

It is recommended that an Advanced Modem Development Committee be formed with representatives from the telecommunications industry, Department of Education, FCC, and NIST, to determine how the required draft standards can best be implemented. A recommended standard should then be submitted to the FCC for processing within its regulatory charter.

The Department of Education should also consider enacting programs in major cities to ease the transition from TDDs to computers through training programs. Personal computers provide more power for the dollar than TDDs, and they benefit from the continual product improvements and cost advantages that intense industry competition promotes. In general, sensory impaired individuals should be encouraged to use products that serve a large segment of the population as these products become cost effective when compared to special-purpose devices such as the Baudot/ASCII TDDs.

15.0 POTENTIAL PROGRAM SCHEDULE

Figure 6 is a proposed schedule for TDD advanced modem development. To ensure that effective standards are developed, the Department of Education should form the Advanced Modem Development Committee, working with Government agencies and the telecommunications industry to incorporate TDD capability at the earliest stages of digital signal processor based ASCII modem development. The Department of Education should evaluate the need for special ASCII/Baudot modem development programs after 9600 bps and 19,200 bps modems have been developed to accommodate TDDs. These two development programs should be designed to put a system in place to inform the telecommunications industry of the needs of the hearing impaired while recognizing the

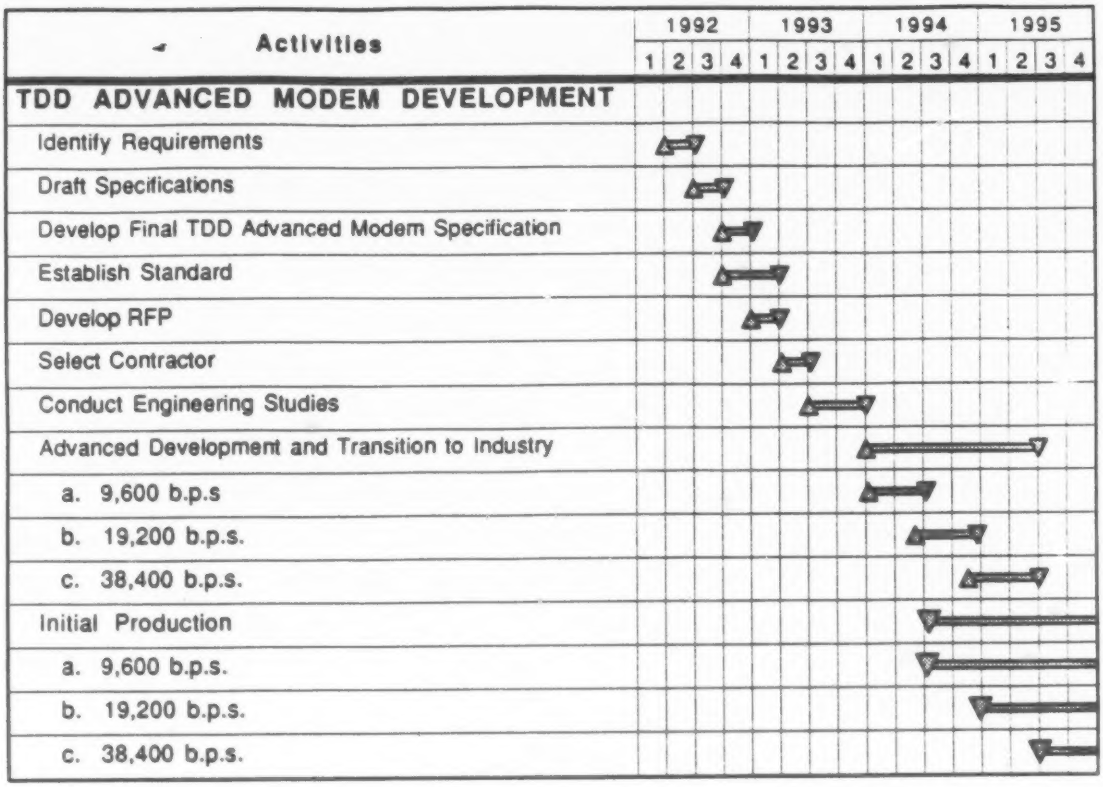


Figure 6. Proposed Schedule for TDD Advanced Modem Development

needs of the telecommunications industry and of the advanced technology modem consumer.

By eliminating the need for new Department of Education modem development projects every two or three years, the TDD Advanced Modem Development Program would reduce future costs to the Department of Education by making this development the responsibility of the telecommunications industry. However, the Department of Education should provide the telecommunications industry with sufficient guidance to make the transition smooth, efficient, and effective. The hearing-impaired community and the general public would then have convenient access to TDDs and ASCII modems with any modem bought off-the-shelf, greatly improving telecommunications access for persons with hearing impairments.

TELECOMMUNICATIONS SYSTEM ACCESS

(Touch Tone Signaling Access)

MARCH 1992

Prepared by

Daniel E. Hinton, Sr., Principal Investigator
and
Dan Morrison, Paolo Basso-Luca and Nancy Davis

SCIENCE APPLICATIONS INTERNATIONAL CORPORATION
3701 N. Fairfax Drive, Suite 1001
Arlington, VA 22203
(703) 351-7755

1.0 SCENARIO

Telecommunications system access such as touch tone signaling access in voice mail applications

2.0 CATEGORY OF IMPAIRMENTS

Persons with hearing impairments.

3.0 TARGET AUDIENCE

Consumers with Hearing Impairments. The consumer with a hearing impairment is unable to use the ever growing number of voice mail and other services which require listening to a message and then responding with a touch tone. If hearing impaired consumers had some means of communicating with such services, they would be able to take advantage of these services. This scenario on telecommunication system access technology such services can be made available to consumers with hearing impairments. The scenario also provides consumers who have hearing impairments with a better understanding of possible opportunities and concepts of system access technology. In addition, this scenario is tied in with the scenario on adaptive modem and TDD access to show how the two work hand-in-hand.

Policy makers, including national representatives, Government department heads, and special interest organizations. Policy makers can use this scenario to better understand the issues related to telecommunications system access for persons with hearing impairments. They may also use it to understand how advanced modem technology combined with touch tone signaling can make it possible to use existing hardware and new software to easily create a new electronic phone service which can be used by both hearing and learning impaired persons.

Researchers and Developers. The R&D community will benefit from this scenario through a better understanding of the phone service needs of persons with sensory impairments. This better understanding of needs will assist researchers and developers in designing functions into future products which will promote a more user-friendly environment for hearing impaired people.

Manufacturers. Manufacturers will benefit through a better understanding of the potential market size and the need for telecommunications access which can be met by adding TDD capability to their modems and new software.

4.0 THE TECHNOLOGY

The Dual-Tone Multifrequency (DTMF) signaling system and the call progress tone standards are the basic technologies associated with telecommunication systems.

Applications associated with telecommunications system access for the hearing impaired are directly related to these technologies.

Standards exist for the Touch-Tone system, which is also known as the Dual-Tone Multifrequency (DTMF) signaling system. DTMF dialing consists of two simultaneously transmitted audio frequencies. On the standard 4 x 3 keypad format, each column and row is associated with a different frequency, as indicated in Figure 1. This method of signaling permits faster dialing for the user and more efficient use of the switching systems. Since the frequencies are in the audio band, they can be transmitted through the telephone network from one user to another after a call has been set up, and also used for data communications.

Frequency	1209 H1	1336 H2	1477 H3
697 L1	1	ABC 2	DEF 3
770 L2	GHI 4	JKL 5	MNO 6
852 L3	PRS 7	TUV 8	WXY 9
941 L4	*	0	#

Figure 1.

An example of data communications using DTMF is the IBM augmented phone services (by IBM Entry Systems Division). This plug-in board & software allows a deaf person to communicate with a hearing person via the IBM computer (without a TDD). The user can type a message on the computer keyboard and the system will send it out over the phone line as synthesized speech. The called person presses keys on the telephone to "spell" the reply (i.e., "BOY" is 269); the software decodes the tones into possible words which the user reads on the computer screen. This provides a technique for basic telecommunication between the hearing impaired and non-TDD equipped individuals.

Other devices that provide use of the DTMF capability for telecommunications access may be exploited. The IBM communicator, described above, is just one example of this application. Also available are the Echo 2000 by Palmetto Technologies Incorporated, and the TE 98 Communicator by Auditory Display Incorporated, which are telecommunication devices for hearing impaired individuals with limited speech skills. The Echo 2000 device works with a standard touch tone telephone, using a two digit number code to represent letters of the alphabet as well as commonly used words and phrases. Incoming messages are displayed on the device's 16 character LCD display. The Echo 2000 and the IBM communicator provide a unique, but limited, interface between the hearing impaired and equipped hearing individuals when the later does not have a TDD.

Telephone systems provide hearing users with feedback on call progress in order to simplify operation and reduce calling errors. This information can be in the form of lights, displays, or ringing, but is most often an audible tone heard on the phone line. These tones are generally referred to as call progress tones, as they indicate what is happening during phone calls. Conditions like busy line, ringing called party, and inoperable number, each have distinctive tone frequencies and cadences (length of time the tone is on or off) based on standards established by AT&T or a telecommunications regulatory agency.

Unfortunately, standards for call progress tones are applied differently in different countries or situations. This report focuses on standards in the United States, but does not preclude recommendations for other systems. Information about most call progress standards is available, and which tones are commonly used can be determined by reviewing references.

In the United States, the call progress standards are defined in AT&T's "Notes on the Network," and for PBX's in the Electronics Industries Association (EIA) RS-464 documentation. Table 1 shows the standard call progress tones as defined in the AT&T document.

Call progress monitoring is currently provided in relatively few TDDs, but could be added at a low cost. The Freedom 415 by Selective Technologies, Inc., TDD and TouchTalk Travelpro by ZiCom Technologies, Inc., have built-in call progress monitoring to indicate dial tone, call ring, busy signal, or voice reception.

5.0 STATEMENT OF THE PROBLEM

A hearing impaired individual is challenged when he/she attempts to access certain parts of the telecommunications network, including: non-TDD equipped individuals, voice mail, automated attendant system, and Public Switches Exchanges (PBX). A brief discussion of these problems follows.

Persons with hearing impairments are challenged by the expanded use of DTMF based applications that make certain tasks easier for those who can hear. These challenges

Table 1. Some Common Call Progress Tone Cadences and Frequencies

DIAL TONE	
Cadence	On, steady
Frequencies	400,425,350 + 440,600 x 120,33 Hz
AUDIBLE RING	
Cadence	2 sec on, 4 sec off. . ., or
	1/3 sec on, 1/3 sec off, 1/3 on, 2 sec off
Frequencies	400,425,440 + 480,420 x 40,450,400 x 25 Hz
BUSY STATION	
Cadence	1/2 sec on, 1/2 sec off. . .
Frequencies	400,425,480 + 620 x 120,450 Hz
REORDER (busy circuits)	
Cadence	1/4 sec on, 1/4 sec off. . ., or
	1/2 sec on, 1 sec off. . .
Frequencies	400,425,480 + 620,600 x 120,450 Hz

fall into four basic areas: all access, PBX and operator intercept, touch tone signaling access, and call progress monitoring.

By the same token, several potential applications of the DTMF system may increase the telephone access of hearing impaired people. For example, many customer support lines are now automated by using DTMF signaling to let the caller indicate his/her needs based on voice questions. This system could easily be adapted for use by the hearing impaired by providing a Baudot detection capability, possibly coupled with an advanced TDD that has a multiline display for the text. If the system detects a modem, either Baudot or ASCII, it will switch into the correct modem mode and provide interaction via the keyboard. The primary component of such a system is the software with modem hardware written specifically for this purpose. While in the Baudot mode, abbreviated menus are required to prevent excessively long menus and phone calls. A current example of this technology is the VCS3500 Versatile Communication System by Microlog Corporation, which offers an add-on TDD interface to automatic call answering systems. The TDD module: 1) provides a menu of options to the hearing impaired callers using a TDD and transfers calls to individual extensions as well as takes personal messages from TDD if a busy or no answer occurs; 2) allows TDD callers to retrieve hundreds of prerecorded informational messages (standard responses to common questions) from the automatic system; 3) interacts with

TDD callers who are requesting forms, publications or specific documents to be mailed to them; and 4) allows in-house users to retrieve, via TDD, personal messages from individuals mailbox which is password protected.

6.0 DEPARTMENT OF EDUCATION'S PRESENT COMMITMENT AND INVESTMENT

A recent U.S. Department of Education Small Business Innovation Research Program Request for Proposal, DOE SBIR RFP #91-024, discussed several areas related to telecommunications system access by individuals with hearing disabilities. The list included: line status monitoring, a modem add-on device (ASCII), an auto-detect switch for FAX, ASCII, and voice calls, and 911 system operator training. The following paragraphs will describe these activities in more detail.

A need exists for an inexpensive device to assist persons with hearing impairments in detecting/identifying important line status signals. The need for this becomes clear when considering a typical situation: (1) dial a number, (2) wait for an answer, (3) if no response, give up! A hearing impaired person does not know if the line was busy, no circuit was available, a wrong number was dialed, the phone was ringing, or a person answered. In other words, unless a TDD reply is received, the caller is left wondering why the call did not go through. This problem is confounded by the dependence of these signals on the telecommunications devices, i.e., PBX, and locality, i.e., country.

Given the current high cost and relative rarity of modems that are both ASCII- and Baudot-capable, adaptation or development of an add-on device to allow standard ASCII modems to communicate with Baudot TDDs is an important issue. Such a solution would provide an easy, affordable way to communicate with a TDD via a personal computer (PC) and a hearing impaired person may also use the same PC to communicate with computer bulletin boards or other services using ASCII. This would eliminate the need for a stand alone TDD and an ASCII modem.

As indicated in the SBIR RFP, a need has developed to discriminate between voice, FAX, ASCII, and Baudot TDD calls. Currently, the technology to automatically switch between FAX, ASCII, and voice exists. The extension of this technology to recognize Baudot TDDs is possible by adapting/developing an add-on controller. This could automate telecommunications tasks which often require human interaction.

Another area of interest is training for 911 system operators. The training material will teach 911 operators how to handle emergency situations involving people who are deaf or hard of hearing.

At the University of Delaware's Rehabilitation Engineering Center on Augmentation Communication, work is underway to define an integrated workstation for deaf individuals. The concept is to bring several applications together in a unified system that offers the advantages of the constituent parts. A key element in this work is to identify the modes of telecommunications that can be effectively used by deaf individuals. Specific areas of interest include: telephone monitoring, touch-tone decoding and voice response.

7.0 ACCESS TO TELECOMMUNICATION SERVICES

Access to many different telecommunications services is desirable for the hearing impaired. These services include PBX services, voice mail services, call forwarding, automated attendant systems, and person (hearing impaired) to person.

8.0 POTENTIAL ACCESS IMPROVEMENTS TECHNOLOGY

Persons with hearing impairments can benefit from enhanced access to telecommunication systems in the following areas which deal with updating or verifying information in a remote computer database: message forwarding systems; financial transaction systems; alarm systems; energy management systems; credit card verification systems; and mail order systems. Persons with hearing impairments will also benefit from enhanced access to cellular telephone media. Current NIDRR projects are detailed in Table 2.

9.0 ADVANCED TECHNOLOGIES

Advanced technologies present several opportunities for improving the access of the hearing impaired to telecommunications services in the areas of all access, PBX and operator intercept, touch tone signaling access, and call progress monitoring.

Voice recognition, which is the subject of a separate scenario entitled "Voice Recognition Systems for Personal and Media Access," could significantly enhance the access of hearing impaired people. The voice recognizer could convert speech into text via TDD display or computer monitor for reading by the caller. To simplify the voice recognition task, a standard regarding synthesized speech could be developed. This would improve access to voice mail systems, automated attendant systems, and other voice based systems.

The Teltone T-310 Telephone Access Unit is an RS-232-C compatible controller for PBX and public telephone lines. The T-310 allows computers, terminals, and other intelligent devices to command such telephone system functions as answering and originating calls, observing call status, sending or receiving DTMF signals, "flashing" the line, and coupling audio sources like speech synthesizers onto the line.

There are three main communication features associated with the T-310. The first feature is DTMF/ASCII conversion. After a telephone connection has been established, the T-310 allows data communication between the called and calling parties via the mechanism of DTMF-to-ASCII and ASCII-to-DTMF conversion. DTMF digits entered at the telephone keypad are converted to their equivalent ASCII characters and forwarded to the computer. In the opposite direction, ASCII characters from the computer are converted to the equivalent DTMF tones and forwarded to the network.

Table 2. NIDRR Projects

Project	Organization	Description
Development of an improved assistive listening system for educational, occupational, and recreational settings.	Oval Window Audio	Assistive Listening System
Rehabilitation engineering center on technological aids for deaf and hearing impaired individuals.	The Lexington Center, Inc.	Assistive Technology
Neural network real-time captioning stenographic unit.	Netrologic	Real-time Captioning
Development of portable, computerized, real-time captioning unit for deaf individuals in courtroom environments.	CADSA, Inc.	Real-time Captioning
Research feasibility of a portable, real time stenographic device for the hearing impaired.	Advanced Technologies Concepts	Real-time Captioning
Adaptation and development of a compact, portable computerized real-time captioning stenographic unit for use in courtrooms where deal lawyers or jurors are accommodated.	Virgus Computer Systems	Real-time Captioning
Development of a portable programmable sound recognition device to promote independence for persons with severe or profound hearing impairments.	Applied Concepts Corporation	Sound-Recognition Device
Feature extraction method for development of a visual telephone for deaf individuals.	University of Delaware/Newark	Telecommunications
Integrated, intelligent, and interactive telecommunications system for hearing/speech impaired person.	Integrated Microcomputer Systems, Inc.	Telecommunications

The second feature is electronic voice. By controlling an audio source such as a speech synthesizer or recorded tape player, the computer can use the T-310's FCC-registered audio interface to establish a one-way voice communications link with the remote party.

The third feature is a live voice. An auxiliary telephone sharing the line with the T-310 may be used for normal conversation after a connection has been established by the T-310.

Voice Answering Systems

There are two types of voice answering systems being used by industry. The first is used for voice mail. When the system answers, the caller is asked to enter a mailbox number and then leave a voice message. This type of system sometimes requires a password. The second type of system is designed to direct the caller to the right type of assistance within an organization. For example, when someone calls an insurance company, the voice answering system would say: if you are on a touch tone phone press 1 for policy renewal; press 2 for policy information; press 3 for operator assistance and so on until all services were covered.

The question is how to provide access to these systems for persons with hearing impairments. The voice mail system is the most difficult since it assumes voice-to-voice contact with no TDD or computer modems. However, if the person with a hearing impairment knows there is a voice answering system, then by observing the TDD light they might know to dial a number for an operator for TDD assistance. For example, the FCC could require that all telephone answering systems use a number such as 1111 that would then dial the TDD relay service or connect to a TDD operator or TDD message that could then direct the person with a hearing impairment to the right location. The alternate 1111 number could also provide the TDD menu for leaving a TDD message. The answering system does not care if it is a TDD tone being recorded or a voice message. The voice mail systems will become more of a barrier to persons with hearing impairments because the voice mail systems are relatively inexpensive and even small offices are adding them. The other alternative is to have a short TDD message, such as TDD press 111, with every menu. However, this can be annoying to hearing people. The TDD message would take approximately 2.46 seconds and could follow the first voice message. Therefore, the hearing person would not think they had reached a modem line. This could be reduced to 1.4 seconds if only TDD 1111 were sent. One could also say that a TDD message will follow and if there is a hearing person please wait.

The automatic menu system could also be handled the same way as the voice mail systems. The only real difference in the two systems is the type of message they are trying to convey and what happens after a selection is made.

The future is telephone relay systems for the hearing impaired. These systems will become more and more automated. Automation will depend on ASCII. We will be moving towards technology which makes the current ASCII-Baudot distinction transparent to the user. Eventually we will switch to machine operated relays based primarily on ASCII. The basic idea is that the hearing impaired will no longer need human intervention to achieve access to telecommunications.

10.0 COST CONSIDERATIONS OF ADVANCED TELECOMMUNICATIONS SYSTEM ACCESS TECHNOLOGY

When the cost of the modifications described above are amortized over all systems to be delivered over the next few years, the cost of the advanced technology will be minimized.

11.0 COST BENEFITS TO PERSONS WITH HEARING IMPAIRMENTS

Development of the technologies required for access to the touch tone signaling system and call progress tones would reduce the cost to hearing impaired persons in several ways. First, phone calls would not be made two or three times to be sure the line is not busy or to determine if a TDD is available at the other end. Automated attendant calls, which give prerecorded greetings and voice prompts that render assistance, would be possible since the hearing impaired could use the automated systems for support. This would open up many new sources of information, i.e., customer service, telephone banking, automated ordering systems, etc.

12.0 PRESENT GOVERNMENT INVOLVEMENT IN TELECOMMUNICATIONS SYSTEM ACCESS

The Department of Education funded two modem projects that are listed in the FY89 NIDRR Program Directory. Many early TDD modem developments were funded by NIDRR and Office of Special Education Programs (OSEPS). One project was entitled "Integrated, Intelligent, and Interactive Telecommunication System for Hearing/Speech Impaired Persons." This Phase II project was awarded to Integrated Microcomputer Systems, Inc., Rockville, Maryland, and featured TTY/TDD and ASCII compatibility, "remote signal control, direct connection to the telephone system, and text-to-speech voice announcer."

A Field-Initiated Research project, entitled "Deaf-Blind Computer Terminal Interface," was awarded to SAIC in Arlington, Virginia, for the development of an acoustical modem interface between the Telebrailler, Microbrailler, TDD, IBM-PC compatibles, and the Commodore 64C.

The Conference Center was awarded a Phase I SBIR from the U.S. Department of Education in October 1991 to build a call progress monitor to allow access to telephone call progress signals.

13.0 TELECOMMUNICATIONS SYSTEM ACCESS TECHNOLOGY TIMELINE

The recent introduction of digital signal processors (DSPs) has made it feasible to include call progress capability in all modems. With DSPs, the inclusion of call progress capability is possible through relatively inexpensive software changes inside the modem, instead of expensive added hardware. If all ASCII digital signal processor based modems are required to provide call progress and Baudot TDD modem capabilities, hearing impaired consumers would have access to line signaling in their Baudot TDD modems and ASCII modems. By 1995, advanced modems will be included with most new computers. If Baudot TDD and call progress compatibility is mandated for digital signal processor based modems, any person or company buying a computer would then have access to all Baudot TDDs and call progress capabilities.

14.0 PROPOSED ROAD MAP FOR TELECOMMUNICATIONS

Competition in the modem industry will ensure that modem transmission rates double every two to three years, up to the theoretical limits of telephone lines and that data compression will push modem's efficient transfer rates well beyond 25,000 bps. Modem manufacturers have the economic incentive to use coding techniques, data compression, and other technologies to push beyond the theoretical limits. An initiative by the Department of Education would ensure that these modems will support ASCII Baudot TDD operations. With research and development action over the next 3-5 years, TDD-capable modems and call progress monitoring for use by hearing impaired people and therefore achieve parity with modems designed for the general public by the mid-1990's. If all new digital signal processing modems sold in the U.S. after 1995 are required to be TDD-capable and provide call progress monitoring, businesses and consumers would, for the first time, be able to buy a single-unit ASCII/Baudot modem at competitive prices, with the features and transmission rates available to all modem users. The cost of Baudot and ASCII TDD modems would be amortized over many thousands of units with the benefits of improved communication access for the hearing impaired and greater access to the hearing impaired by the general population. Three to five years after regulatory or legislation is enacted requiring all advanced technology modems to be TDD-capable and provide call progress monitoring, several million modems would have been replaced with Baudot TDD-capable modems and TDD access would be almost universal.

15.0 POTENTIAL PROGRAM SCHEDULE FOR TELECOMMUNICATIONS SYSTEM ACCESS

Figure 2 is a proposed schedule for telecommunications systems access technology development to meet the needs of persons with hearing impairments.

Activities	1992	1993	1994	1995	1996
Identify Requirements	X				
Draft Specifications	X				
Establish Standard		X			
Conduct Engineering Studies			X	X	X

Figure 2. Proposed Schedule for Telecommunications System Access Technology Development

In particular, the Department of Education needs to continue to identify specific needs and applications for telecommunications systems access technology to meet the need of persons with hearing impairments.

VOICE RECOGNITION SYSTEMS FOR PERSONAL AND MEDIA ACCESS

MARCH 1992

Prepared by

Daniel E. Hinton, Sr., Principal Investigator
and
Charles Connolly and Nancy Davis

SCIENCE APPLICATIONS INTERNATIONAL CORPORATION
3701 N. Fairfax Drive, Suite 1001
Arlington, VA 22203
(703) 351-7755

1.0 SCENARIO

Voice Recognition Systems for Personal and Media Access

2.0 CATEGORY OF IMPAIRMENTS

Persons with Hearing Impairments.

3.0 TARGET AUDIENCE

Consumers with Hearing Impairments. Persons with sensory impairments will benefit from enhanced access to media information and telecommunications services. This scenario on advanced voice recognition technology will provide a means to disseminate information to consumers with hearing impairments. It will provide consumers with hearing impairments a better understanding of advanced technology voice recognition capabilities beyond the slow (i.e., 30 to 60 word per minute (wpm)) single utterance voice recognition systems that are in use today.

Policy makers, including national representatives, Government department heads, and special interest organizations. Policy makers can use this scenario to better understand the issues related to spoken media access for persons with hearing impairments. They may also use it as a point of departure in the process to establish research objectives and setting priorities for exploiting future voice recognition technology to meet the media access needs of persons with hearing impairments.

Researchers and Developers. The R&D community will benefit from this scenario through a better understanding of where voice recognition technology is and where it will be in three to five years with respect to meet the communication needs of persons with sensory impairments. A better understanding of voice recognition requirements assists researchers and developers in designing voice recognition functions in future products and promoting an environment in which the needs of persons with hearing impairments are met.

Manufacturers. Manufacturers will benefit through a better understanding of the potential market size and the existing Federal Government requirements for media access which can be met by adding voice recognition capabilities to their products.

4.0 THE TECHNOLOGY

Voice recognition systems technology encompasses everything from simple user-trained computer software and electronic circuits capable of recognizing a few single utterances user adaptable continuous speech speaker-independent systems capable of 1,000 to over 20,000 words. Although the speaker-dependent systems have been on the market for over 10 years, the advanced technology speaker-adaptable continuous voice recognition systems are just beginning to make their appearance, and the speaker- independent continuous voice recognition systems are in research and development. These systems are expected to be available within 3 to 5 years for specific applications such as medical transcription.

The early voice recognition technology was template matching systems made up of the building blocks shown in Figure 1. The systems include a microphone, amplifier, analog to digital converter and a recognition algorithm capable of extracting the information necessary to identify a single spoken word based on a word template stored in computer memory. These systems matched only a few words (i.e., 16-100) at most. The limitation was due to the processing power of the computers and the available memory (i.e., each word required several hundred bytes of memory). These machines matured into more advanced models that recognize several thousand words (i.e., 1,000-10,000) provided a pause is inserted between words. Most of the machines on the market today fall in this category. Table 2, in Section 10, is a representative list of advanced voice recognition systems and their costs on the market today.

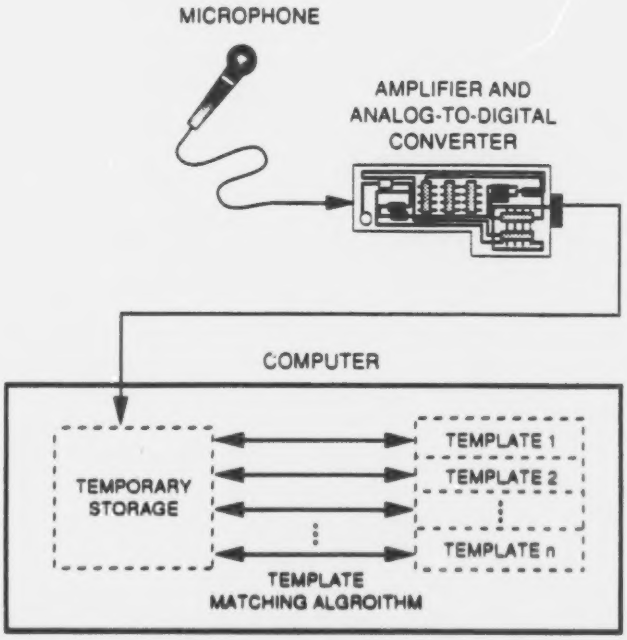

MICROPHONE
AMPLIFIER AND ANALOG-TO-DIGITAL CONVERTER
COMPUTER
TEMPORARY STORAGE
TEMPLATE 1
TEMPLATE 2
TEMPLATE n
TEMPLATE MATCHING ALGORITHM

Figure 1. A Template Matching Speech Recognition System

The advanced technology voice recognition systems are using new computer digital signal processor boards, statistical software and advanced acoustic microphone technologies to achieve speaker adaptable, speaker independent continuous speech recognition systems that can recognize words, and form them into sentences in real time. One system under development from Dragon Systems Inc. uses an IBM PC 386 or 486 with a digital signal processor board and advanced statistical software to recognize over 10,000 words of single user adaptable natural speech. Table 1, in Section 9, is a list of organizations developing advanced technology voice recognition systems. These systems are being evaluated by the

Defense Advanced Research Project Agency (DARPA) and their progress is being monitored through an annual Speech and Natural Language Workshop. This paper describes the current state of technology, where it is expected to be in three to five years, and how it could be applied to meet the needs of persons with hearing impairments.

To advance the technology beyond single word recognition requires a completely new paradigm for voice recognition. The DARPA Information Science and Technology Office initiated a workshop on voice recognition technology in 1989 to explore new methods for speech and natural language processing. Their approach was to treat the subject of voice recognition as a speech and natural language process. This process recognizes that before speech can be recognized, the entire speech and language process must be understood and an interchange of information must take place between various scientific disciplines. Prior to the exchange of information, a common set of definitions on what speech and natural language is, what features are relevant, and the goals of each discipline must be determined.

Voice recognition systems are divided into two classes: feature based and speech trained. Feature based systems explore spoken words to determine characteristics of the vectors (i.e., composition of the words and spectral content) and to determine what common invariant behavior they have. From these vectors, characteristics rules are formulated which can then be applied to the recognition process. For many years, the speech recognition community has been trying to perfect a feature based system because, although probably 10 to 15 years in the future, they offer the most versatile features. Only speech trained systems will be discussed in this scenario. In the speech trained systems, speech is used to train the system automatically. There are currently three methods for accomplishing the training: template matching, statistical modeling, and neural nets. The descriptions of template matching and statistical modeling that follow are based on "Speech Tutorial," by Edward Neuberg, which appears in DARPA's "Proceedings: Speech and National Language workshop," February 1989 (distributed by Morgan Kaufmann Publishers, Inc., 2929 Campus Drive, San Mateo, CA 94403, ISBN# 1-55860-073-6).

Template Matching Systems

The simplest template matching systems recognize 1 to 100 words. These systems store a speaker's words as a series of vectors called a template. When the speaker says the word again the system vectorizes its amplitudes, frequencies, and duration and compares it with the persons speaking templates stored in memory. The template with the best match is declared the winner and the word with the best match is selected. Implicit in these systems are:

- the ability to find the beginning and ending of each word
- a quantitative way of comparing two utterances (i.e., scoring algorithm).

There are two algorithms that have been applied to these problems. The first is Dynamic Time Warp (DTW) that allows one to compare sequences of different length by stretching one sequence to the same length as the other. The second is a more sophisticated version of DTW called Level Building that allows template matching to be used on connected speech, in which the beginning and ending of a spoken word cannot be found. This algorithm is a brute force method that applies all possible time warps to every utterance, compares them with all templates in the data base and selects the one that scores the highest. The memory size and processing speed needed to achieve large vocabulary systems based on this approach are generally considered prohibitive beyond several hundred words. This also considers that not only must the words be matched but the task domain restrictions must be considered (i.e., grammar, semantics, subject matter, and pragmatics).

Template systems are generally applied to single speaker voice recognition systems, although by training the system using several speakers, some degree of speaker independence can be achieved.

Statistical Modeling Systems

Statistical modeling systems have been developed because sound spectrum sequence analysis is too complicated at this time to determine all of the rules necessary to identify certain utterances as words. Template matching is impractical because the variability of pronunciation is too great, and phoneme templates have just not worked. To overcome the limitations of these systems, statistical models have been developed to extract the statistical properties of sound. These models are based on extremely simple machine states. The form of the machine is assumed, and then its parameters are statistically estimated using a large amount of speech data. Currently the Hidden Markov Model (HMM) has been the most widely used statistical model. For this scenario, the HMM work is considered the most advanced and the systems presented use this model. To better understand the systems, a general description of how the model works is needed without presenting a rigorous mathematical proof.

The Markov model is a state machine made up of a finite number of discrete statistical states. Each state represents a piece of a word (or a word, sentence, etc., in a hierarchy. Basically, the machine has two parts. The first part determines what state the machine is in. For the purposes of this discussion, 5 states are assumed although each model of the HMM can have any number of states depending on the implementation. The state machine has a centisecond clock and a probabilistic change rule: the machine starts at state 1, and every centisecond it applies the rules to determine what the next state will be. The probability of transitioning to any given state depends only on the present state (which makes the model a Markov model). The machine continues this process until it reaches the final state: state 5.

The second part of the machine actually determines what vector (that is, sound) will come out of the machine between the most recent centisecond clock tick and the following one. the probability of a given sound coming out depends entirely on what state the first part of the machine is in. That state's vectors have a given statistical average and variance. Note that, so far, the Markov model is actually modeling the person speaking, not the computer "listening." A particular sequence of states sequence of sounds (vectors) a particular word. Somewhat different sequences of sounds (vectors) may result from the same sequence of states because no one says the same word exactly the same way every time.

What makes the Markov model "hidden" is when it is applied to speech recognition. Given a sequence of sounds (vectors), the model includes enough information to determine the probability that those sounds correspond to a given sequence of states, representing a particular word. However, there is no way to "see" which sequence of states produced the sounds; that information is hidden. All that can be done is to find the probabilities that various sequences of sates (words) produced, the observed utterance, then pick the one with the highest probability.

For example, suppose one assigns probabilities to 5 transitions of part one of the machine and to the 5 means and variances of part two of the machine. Next one collects a large number of tokens of some word and calls them observations. There is a statistical technique for calculating the probability that the statistical parameters are correct, given the observations. Then there is an algorithm that allows one to choose the best set of statistical parameters given a set of observations. Applying this algorithm repeatedly allows one to "climb" to a set of parameters for the machine that is most likely to be correct for the observations within the data set. This set of parameters--transition probabilities and means and variances--is the statistical model for the word(s) collected and trained.

To train the system, one tokenizes a set of words that are used as the models to apply to the HMM and form a statistical model of the speech to be processed. For each word the machine "hears," the model calculates the probability that incoming word was produced by the state sequence representing each word in the model; if there are 50 words then there are 50 probability estimates and the highest probability estimate is chosen as the word.

The advantage of the HMM is that one need not segment the training collection or the incoming word. The disadvantages are that one must assume that every word, of whatever length, has the same number of states and one can never know what the states mean. However, this is not a mathematical problem, only an intuitive problem for a researcher.

In all the above discussions, one must keep in mind that this process is continuously variable. Some of the sources of variability that plague Voice Recognition Systems include:

- size, sex, age
- dialect
- loudness, speech rate
- health, emotion
- coarticulation

- channel
- noise.

Neural Net Systems

Neural networks have been used for small (1-100 word) vocabulary speaker-independent applications. However as the number of words increase, the training time and complexity of the networks increases and the system performance decreases. In addition, there is no known efficient search technique for finding the best scoring segmentation in continuous speech using neural networks alone. To overcome these problems, hybrid systems are being employed that take advantage of the HMM search capabilities to find the "N" best matches and then employ segmental neural networks (SNNs) computational advantages to evaluate those matches. Figure 2 illustrates a hybrid system developed by BBN Systems, and Technology, Inc. and Northeastern University in Boston for a continuous speech recognition system using segmental neural networks. The work was funded by DARPA.

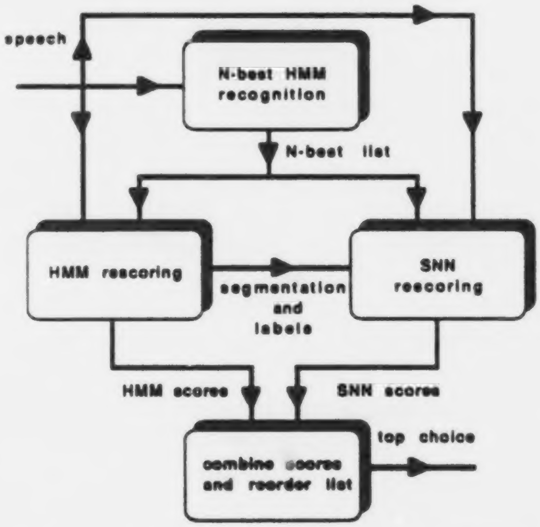


Figure 2. N-Best Rescoring System Using the SNN Score. (p. 251, Feb book)

With the Hybrid SNN/HMM system, the HMM approach is used to obtain the N-Best word list (i.e., highest probability matches). Next the HMM and SNN system are run in parallel to evaluate those word matches. The HMM provides intermediate results to the SNN so the SNN can score the proposed word choices. Finally a combined score is obtained. The word with the best combined score is the one chosen. Using this system the worst case word recognition is as good as an HMM system alone. This hybrid HMM/neural network approach seems to offer an improvement in voice recognition accuracy. It is estimated that this technology is three to five years from maturity in the laboratory.

5.0 STATEMENT OF THE PROBLEM

Advanced voice recognition systems are being developed to meet the needs of Government and the American consumer, for high quality data entry (transcription) and machine control. At this time, access to single word voice recognition technology for persons with sensory and physical impairments is being addressed by Government and research institutions (i.e., management, researchers, and marketers). Persons with hearing impairments need a high quality speaker independent continuous speech recognition system to provide interpreter services for face-to-face, public address, mass media and telephone media access. Because the single word voice recognition systems require a pause between words, they are limited to approximately 60 words per minute maximum and are not practical for use as an interpreter system for persons with hearing impairments. The average speaker speaks at a rate of 150 to 200 words per minute without pauses between words. For services such as closed captioning for television news, the rate can be as high as 270 words per minute. Clearly, to meet the requirements of persons with hearing impairments for natural voice processing, advanced technology voice recognition systems are required that are speaker independent and can translate speech to text in real time.

6.0 DEPARTMENT OF EDUCATION'S PRESENT COMMITMENT AND INVESTMENT IN VOICE RECOGNITION SYSTEMS

Presently the Department of Education has no investment in voice recognition systems. However many of the goals and objectives of the Department of Education could be met with a high quality user independent voice recognition system. Examples of the goals and objectives of the Department of Education are as follows:

- In the Department of Education's "Small Business Innovative Research (SBIR) Program Phase I Request for Proposal," issued January 11, 1991, research topics related to telephone and media access included:
    - Adaptation or development of an add-on controller which will enable telephone switching devices to automatically recognize incoming Baudot TDD calls and switch them to the correct device--a capability which is currently available for FAX, ASCII, and voice calls that come in on the same telephone line. Voice recognition systems could be adapted to do this function.
    - Adaptation or development of an inexpensive device which will assist persons with hearing impairments to detect and/or identify some of the most important telephone line status sounds in a particular locality. Voice recognition systems could be trained to perform this function and provide other functions to augment telephone access to the general public.

- The findings of the National Workshop on Rehabilitation Technology, a cooperative effort of the Electronic Industries Foundation (EIF) and the Department of Education's National Institute on Disability and Rehabilitation Research (NIDRR), indicated a need for research to develop computerized techniques to facilitate the use of telephone systems and broadcast media by deaf, hard-of-hearing, and visually impaired/hearing impaired persons, including voice/Baudot TDD interfaces. The addition of voice recognition to telephone communications and media communications could significantly impact this recommendation.

- The *Federal Register*, December 4, 1990, states the Final Funding Priorities for the NIDRR for fiscal years 1991-1992. These priorities include "creating more accessible communication environments for (the deaf and hard of hearing) population." One of the stated approaches to meeting that goal is to "conduct at least one national study of the state of the art to identify current knowledge and recommend future research."

The program, titled "Examining Advanced Technologies for Benefits to Persons with Sensory Impairments," conducted by the U.S. Department of Education who developed this scenario, represents one such study. The Panel of Experts for this program included nationally known experts in technology and persons with sensory impairments. When the Panel met February 7-8, 1991, there was a consensus that user independent continuous voice recognition is a high priority need for persons with hearing impairments. The reason cited for its inclusion was that a voice recognition capability would have a substantial impact on telephone and telecommunications media access for persons with hearing impairments as well as a significant impact on personal communications with the general public.

7.0 ACCESS TO VOICE RECOGNITION TECHNOLOGY FOR INFORMATION AND COMMUNICATION MEDIA ACCESS

There are many federal, state, and local laws which influence telecommunications for hearing impaired people, just as these laws influence telecommunications for the general population. However, the most important single law related to telecommunications for hearing impaired people is Public Law 101-336, enacted July 26, 1990. Better known as the Americans with Disabilities Act (ADA), this law has broad implications for all disabled Americans and establishes the objective of providing access to persons with disabilities to electronic and physical facilities and media.

The other law that impacts technology for persons with hearing impairments is Public Law 100-407-Aug. 19, 1988, titled "Technology-Related Assistance for Individuals with Disabilities Act of 1988," which established a comprehensive program to provide for technology access to persons with disabilities. In this law, assistive technology devices are defined:

> "Assistive technology devices means any item, piece of equipment, or product system, whether acquired commercially off the shelf, modified, or customized, that is used to increase, maintain. or improve functional capabilities of individuals with disabilities."

Voice recognition technology clearly meets this definition for persons with hearing impairments and should be exploited to increase the opportunities for persons with hearing impairments to obtain access to voice media and individual services. Within the Findings and Purpose of this law, voice recognition technology can provide hearing impaired persons with hearing impairments with:

- greater control over their own lives;
- participation in and greater contribution to activities in their home, school, work environments, and communities;
- greater interaction with hearing individuals; and
- other benefits from opportunities that are taken for granted by individuals who do not have sensory handicaps.

8.0 POTENTIAL ACCESS IMPROVEMENTS WITH ADVANCED VOICE RECOGNITION TECHNOLOGY

Potential telecommunications and media access improvements for persons with hearing impairments will be significantly impacted with the advent of speaker independent continuous voice recognition systems as follows:

- *Face-to-face communications with the general public*. This is the most exciting application of voice recognition. The vast majority of hearing people do not know sign language. Most persons with hearing impairments have a difficult time with lip reading. User independent voice recognition that was only 70 to 80 percent accurate, given the hearing impaired persons ability to process the speech and choose the correct word in a sentence, would be a significant step forward in face-to-face communications. Given a large vocabulary voice recognition system (10,000-30,000 words or more) in an environment that suggests some context (classroom, laboratory, office, home, etc.) an excellent exchange of information via voice to text might be expected.
- *Telecommunications media access*. Systems with large vocabulary voice recognition could be integrated into the telephone system Telecommunications Devices for the Deaf (TDD) to replace existing relay systems between hearing individuals and persons with hearing impairments. In a phone system, the computing power could be significantly increased over individual

computer based systems. Several hundred thousand word voice recognition systems could be available on-line. The system could work as follows:

-- A hearing person could make a call to a hearing impaired person or vice versa.

-- Upon receiving or initiating a call, the hearing impaired person could press the pound ("#") key to activate the system.

-- The system could provide TDD-type output to the hearing impaired person and voice output to the hearing person (i.e., either the hearing impaired person's voice or synthetic speech, depending on what service has been preprogrammed).

-- The hearing person could provide voice feedback on the translation if requested, with the asterisk ("*") key on the telephone touch tone keyboard.

-- To achieve this, noise suppression systems and small bandwidth voice recognition systems (i.e., 3,500 to 5,000 Hertz) bandwidth would be required. It is estimated that with a concerted effort such a system could be fielded within 5 years for vocabularies of up to 1,000 words. Signition, Inc. Hearing Research Laboratory is working on the noise suppression circuits that may make this application viable.

- *Communications media access (TV, recordings, radio, public address systems).* One of the most exciting applications of voice recognition technology is in the area of television closed captions for the hearing impaired. Today, real time closed captions are generated by a stenographer using special equipment. Only the fastest stenographers can keep up with a real time news broadcast. These stenographers are paid $60,000 to $80,000 per year. Voice recognition technology in its present state could provide closed captions using much less skilled individuals with salaries of $30,000 to $40,000 per year. This would allow the cost of closed caption programming to be reduced 30 to 40 percent, given that administrative and editing costs would not be reduced.

A voice recognition closed caption system would consist of a person to repeat the broadcast (i.e., transcriber), a voice recognition transcription computer system, and the existing closed caption equipment. The voice recognition system could be trained to a set of voice transcribers to increase the recognition rate to 98% accuracy or higher. As improvements are made to the voice recognition systems, the systems could directly transcribe the commentator's voice or the voice of any persons being interviewed. This system would overcome several problems with laboratory systems such as noise, accents, and repeats. The transcriber could also add punctuation by simply pressing a comma, period, question mark, or semicolon key. Several systems presently under development by DARPA could be adapted to this purpose. To accomplish this would require several steps:

-- *The development of a word recognition database for the type of programs to be real time captioned.* This database of words could be created based on existing databases at captioning centers such as WGBH in Boston. An example would be news programs. A person could establish a word set based on one to two years of a news program's previous closed caption files. One could capture the closed caption files as an ASCII computer file and analyze the files for word and grammar content. A database could then be created for the words and search strategies could be devised for the most common words. In a given situation, most writers tend to use 1,000-3,000 words. Individual spoken vocabularies consist of 1,000-1,500 words. The database just described is estimated to be approximately 20,000 words. Organizations such as the National Captioning Institute or GBH in Boston could easily quantify the size of such a database from existing closed caption files or video recordings.

-- *Develop a training methodology for voice transcribers.* Methods of training voice transcribers must be developed to optimize the voice recognition training time. Vocabularies for the training program need to be developed and training strategies perfected. It is estimated that as little as 40 hours of training time for the voice transcriber and an additional 40 hours of voice recognition system training could result in a system with 99% or better voice transcription accuracy for real time captioning using one of the systems being developed for DARPA.

-- *Develop an interface between the voice recognition system and the existing closed caption hardware.* This will be the easiest part of the program with the least risk. The voice recognition systems are on SUN workstations or IBM PC 486 compatible machines with RS-232C interfaces. All that is required is the development of interface software that provides the voice transcription to the closed caption hardware at a specified rate and in the closed caption format for three line rollup (i.e., the predominant form of live caption).

- *Interpreter services (education, business).* Providing interpreter services is a harder problem to solve than closed caption voice transcription since these services are used in a high ambient noise environment. Television closed caption voice transcribers can be placed in a soundproof room. For meetings and auditorium type speaking, the voice transcriber interpreter could also be in a soundproof room. The transcription could be displayed on a screen or moving display. Another approach would be to use a remote interpreter via a telephone hookup. The voice would be transmitted to a studio, transcribed, passed back by modem over the phone, and finally displayed. With the advent of cellular telephones, this would offer a simple way to interpret for even small conferences in large cities, reducing the travel time and cost of interpreter services. For Government and business it may offer an inexpensive way to conduct telephone relay or provide local interpreter access for persons with hearing impairments.

9.0 ADVANCED VOICE RECOGNITION TECHNOLOGIES

DARPA sponsors much of the most advanced unclassified research in speech recognition in the U.S. (see Section 12). The following advanced voice recognition systems are representative of the state-of-the-art in speech recognition and natural language processing and are presented in DARPA's "Proceedings: Speech and Natural Language Workshop," June 1990, (distributed by Morgan Kaufmann Publishers, Inc., 2929 Campus Drive, San Mateo, CA 94403, ISBN# 1-55860-157-0) and "Proceedings: Speech and Natural Language Workshop," February 1991, (also distributed by Morgan Kaufmann Publishers, Inc., ISBN# 1-55860-207-0). These speech recognition systems represent significant steps forward in user dependent and continuous user-independent voice recognition systems that can be applied to the needs of persons with hearing impairments and physical handicaps. The synopses of the articles presented here illustrate the advanced technology presently being applied to voice recognition within the research and development community.

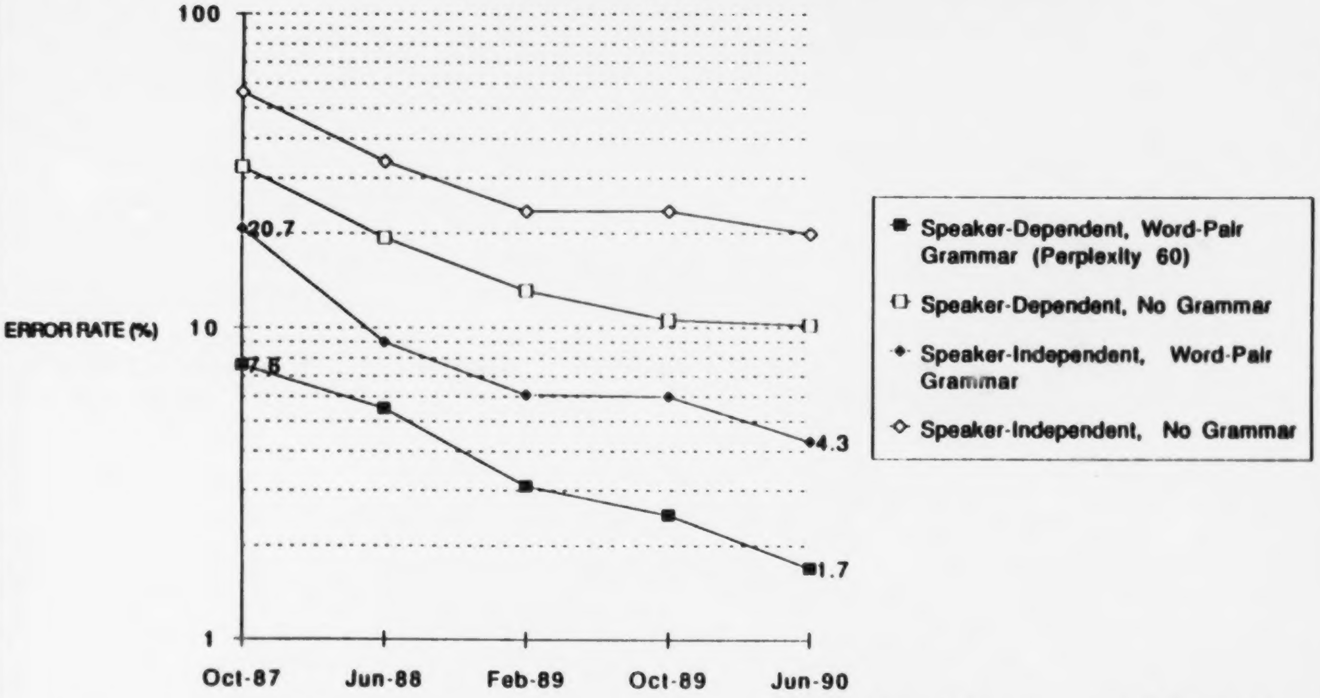
This section focuses on voice recognition and the translation of continuous speech into text. It is assumed that for persons with hearing impairments that most of the natural language processing will be performed by the person and not the machine. However, the machine can improve the output to the person with hearing impairment by having the capability to examine the language structure to check for misinterpreted words and phrases. Therefore, some of the systems go beyond voice recognition systems and are speech and natural language processing systems.

- *A Snapshot of Two DARPA Speech and Natural Language Programs.* DARPA's Spoken Language program has two major components: large vocabulary speech recognition, which has many applications, and spoken language understanding, aimed at interactive problem solving. Both deal with spontaneous, goal-directed, natural language speech and both also aim for real-time, speaker-independent or speaker-adaptive operation. The program also includes basic research to fuel the next generation of advances. Progress on continuous speech recognition is presented in Figure 3.

  Performance evaluation for speech recognition is currently being conducted using the Resource Management (RM) corpus, which consists of read queries and commands, and the Air Travel Information System (ATIS)

PROGRESS ON CONTINUOUS SPEECH RECOGNITION

RESOURCE MANAGEMENT CORPUS, READ SPEECH
1000-WORD VOCABULARY, PERPLEXITY 60



EVALUATION ——> PROGRESS

Figure 3. Progress on Continuous Speech Recognition

corpus, which consists of spontaneous queries and commands. Plans are underway to expand the ATIS corpus and to replace the RM corpus with a more challenging one.

Performance evaluation for speech understanding is being conducted with the ATIS corpus, collected from subjects interacting with a simulated (wizard-based) understanding system that contains certain data from the Official Airline Guide (OAG).

In addition, several groups are also developing spoken language technology demonstration applications. The most advanced of these is MIT's Voyager system, which provides navigational assistance for Cambridge, MA.

Groups currently being funded include BBN, Brown, BU, CMU, Dragon, Lincoln, MIT, SRI, TI, and UNISYS. The program is greatly enriched by the voluntary participation of AT&T in the periodic performance evaluations.

Charles L. Wayne, DARPA/ISTO
3701 N. Fairfax Dr., Arlington, VA 22203; (703) 696-2259

- *Toward a Real-Time Spoken Language System Using Commercial Hardware.* BBN Systems and Technologies, Inc. describes methods and hardware used to produce a real-time demonstration of an integrated Spoken Language System. Algorithms that reduce the computation needed to compute the N-Best sentence hypotheses are detailed. A fully-connected first-order statistical class grammar is used to avoid grammar coverage problems. The speech-search algorithm, which is implemented on a circuit board using an Intel i860 chip, plugs into the VME bus of a SUN4. This controls the system and contains the natural language system and application back end.

  To eliminate machine dependence, all code was written in C. With a combination of algorithms, code optimization and faster hardware, they were able to speed up the N-Best computations by a factor of 20,000 and achieve better than real-time continuous voice recognition.

Steve Austin
BBN Systems and Technologies, Inc.
10 Moulton St., Cambridge, MA 02138; (617) 873-2000

- *Developing an Evaluation Methodology for Spoken Language Systems.* Only recently has progress been made in agreeing on a methodology for comparative evaluation of natural language and speech understanding. BBN Systems and Technologies, Inc. presents the DARPA/NIST, an evaluation methodology. In their paper they detail the process that was followed to create a meaningful spoken language system evaluation mechanism, describe the current mechanism, and then present directions for future development of speech recognition and natural language. This effort formalizes the methodology for measuring natural languages and speech understanding systems in order to measure progress.

Madeleine Bates
BBN Systems and Technologies Corporation
10 Moulton St., Cambridge, MA 02138; (617) 873-2000

- *The Dragon Continuous Speech Recognition System: A Real-Time Implementation.* Dragon Systems, Inc. presents a 1000-word continuous speech recognition (CSR) system which operates on a personal computer in near real time. This system, which is designed for large vocabulary natural language tasks, uses Hidden Markov models (HMM) and includes acoustic, phonetic and linguistic sources to achieve high recognition performance. By using advanced algorithms with software optimizations, computation requirements have been reduced by a factor of 30, with little loss in performance. When using a 386-based PC, the recognizer was clocked at 2.8 times real time, 1.5 times using a faster 486-based PC, and 1.1 times when using a 29K-based digital signal processor add-on board. User dependent voice recognition results on a single speaker, using 1000 test utterances totaling 8571 words (847 different words), were 293 word errors (3.4% word error rate).

Paul Bamberg
Dragon Systems, Inc.
Newton, MA 02158; (617) 965-5200

- *Recent Progress on the VOYAGER System.* The Massachusetts Institute of Technology's (MIT) VOYAGER speech recognition system is an urban exploration system providing the user with help locating various sites in the Cambridge, Massachusetts area (i.e., banks, restaurants, and post offices) using either voice input or typed input.

  Two main developments are a tighter integration of the speech and natural language components and the implementation of a pipelined hardware which leads to a speed-up in processing time to approximately 5 times real time. There have also been improvements made to word-pair grammar, pronunciation networks, and the back-end capabilities.

  As of June 1990, the VOYAGER system correctly recognized almost 30% of 4361 sentences, totaling about 35,000 words (601 different words). That figure is now on the order of 50% with word recognition error rates of 6-7%. For comparison, Dragon's sentence error rate was 19.5%. Note that

VOYAGER is speaker-independent; Dragon's system was tested with only one speaker.

Victor Zue, Rm NE43-601
Spoken Language Systems Group, Lab for Computer Science
MIT, Cambridge, MA 02139; (617) 253-1000

- *Recent Progress on the SUMMIT System.* MIT's SUMMIT system is a speaker-independent, continuous speech recognition component of VOYAGER. It has a vocabulary of up to 1000 words with perplexities of up to 73. The difference between this system and other HMM approaches is its use of auditory models and selected acoustic measurements and its segmental framework and use of pronunciation networks.

  MIT has integrated SUMMIT with a natural language system. They have also changed the normalization procedure to make it more responsive for recording spontaneous speech. This system is being used as a test bed for new algorithms and hardware to evaluate speech and natural language systems.

Victor Zue, Rm NE43-601
Spoken Language Systems Group, Lab for Computer Science
MIT, Cambridge, MA 02139; (617) 253-1000

- *Research and Development for Spoken Language Systems.* MIT is developing a spoken language system that will demonstrate usefulness of voice input for interactive problem solving. The system, which combines SUMMIT, a segment-based speech recognition system, and TINA, a probabilistic natural language system, accepts continuous speech and handles multiple speakers. Accomplishments are:

  -- Improved performance and expanded capabilities of the VOYAGER urban exploration and navigation system.

  -- Developed a mechanism for generating tasks automatically in the VOYAGER framework in order to promote interactive problem solving by users which enabled MIT to collect spontaneous speech from users in a goal-directed mode.

  -- Performed acoustic and linguistic analysis on nearly 3,000 sentences, contrasting read and spontaneous speech.

  -- Developed an initial version of ATIS, collected pilot data, and participated in the first round of common evaluation.

MIT's future plans include:

- Improving speech recognition performance by incorporating context-dependency in phonetic modeling.
- Fully integrating TINA and SUMMIT to exploit speech and natural language symbiosis.
- Continue increasing and improving the knowledge base of VOYAGER, in order to generate correct and natural responses.
- Collecting additional speech and text data during actual problem solving for system development and evaluation, and continuing to evaluate the performance of VOYAGER.
- Continuing hardware development, so that the system will be able to run in near real-time.

Victor W. Zue, RM NE43-601
Spoken Language Systems Group, Lab for Computer Science
MIT, Cambridge, MA 02139; (617) 253-1000

- *Speech Research at Carnegie Mellon*. Carnegie Mellon University is conducting research to effectively integrate speech into the computer interface. They are trying to eliminate fundamental limitations of current speech recognition technology. Current areas of research are:
  - Improved Recognition Techniques. Developing a 5000-word, speaker independent, connected speech recognition system.
  - Fluent Human/Machine Interfaces. Studying the utility of speech in day-to-day interactive tasks.
  - Acoustical and Environmental Robustness. They have developed algorithms dealing with several classes of variability in speech signal. This includes noise-subtraction algorithms based on traditional approaches that substantially eliminate stationary noise interference.
  - Understanding Spontaneous Spoken Language. By using sophisticated parsing techniques dealing with ill-formed speech, they have moved beyond small languages and rigid syntax in situations where the user cannot learn a restricted command language.

-- Dialog Modeling. By applying the work begun with the MINDS system to more domains they have successfully shown that similar dialog-level constraints can be applied to recognition.

Raj Reddy
School of Computer Science
Carnegie Mellon Univ., Pittsburgh, PA 15213; (412) 268-2565

- *Speech Representation and Speech Understanding*. Speech Systems, Inc. (SSI) is conducting research to encode speech into segments which retain information necessary for accurate and continuous speech recognition, but which are more efficient to deal with than the usual encoding of short frames of speech. They are using a multi-stage decision tree encoder with linear combinations of features at the decision nodes.

They are also working on proving that application knowledge can be efficiently applied to these codes in order to produce accurate transcriptions. Their objective is to produce results which can be employed in a commercial system. Accomplishments that have been achieved are:

-- Showed that by segmenting and coding speech using SSI's phonetic encoding there was a significant improvement of both speed and accuracy for a system using Markov modeling.

-- Reduced utterance error rate by 25.7% with a 40% increase in speed by reducing the number of ways words were spelled in a dictionary and by re-defining the phonetic classes.

-- Word error in decoding phonetic codes into words was further decreased, typically 20%, using a penalty that reduced the erroneous insertion of small words.

-- Further increased speed of recognition by a factor of two by using a more efficient structure in the decoding software.

-- Modified software to provide access to transcriptions other than the best guess (e.g., the second through tenth best guesses) to aid the user in making corrections.

-- Also modified software to give application developers access to semantic knowledge inherent in structure of language model used in the recognition process. For example, various names that a radiologist called a tumor seen in an x-ray (mass, density, tumor, etc.) would all be labeled "tumor."

William S. Meisel
Speech Systems, Inc.
18356 Oxnard Street, Tarzana, CA 91356; (818) 881-0885

- *A Real-Time Spoken-Language System for Interactive Problem-Solving*. SRI is developing a system which improves complex problem-solving through the use of interactive spoken language in concert with other media. To do this requires real-time performance, large vocabulary, high semantic accuracy and habitability. Their system is being developed in the air travel planning domain using two research and development lines: focusing on a spoken language system kernel for database query and on full interactive systems. Word recognition rates are exceeding 98% accuracy.

Patti Price
SRI International
333 Ravenswood Ave., Menlo Park, CA 94025; (415) 336-6200

The following is a synopsis of an article, *ESPRIT, The European Strategic Programme for Research and Development in Information Technology*, published in DARPA's "Proceedings: Speech and Natural Language Workshop," February 1991, by Mr. M. Moens of the University of Edinburgh.

Mr. Moens obtained the following information from the Information Package for Proposers, 1989. The European Strategic Programme for Research and Development in Information Technology (ESPRIT) has the following objectives:

- Provide the European industry with basic technologies to meet competitive requirements of the 1990s.
- Promote European industrial cooperation in pre-competitive R&D in information technology (IT).
- Pave the way for internationally accepted standards.

Three major information technology sectors are being addressed by ESPRIT:

- Microelectronics
- Information Processing Systems
- Applications.

ESPRIT is implemented through pre-competitive R&D projects. Topics addressed are described in the ESPRIT Workprogramme, which also reports the strategy, objectives and technical aspects of the program. The Workprogramme is updated on a regular basis in consultation with the European IT community. Requests for proposals are published in the Official Journal of the European Communities.

ESPRIT projects are performed under shared-costs contracts by consortia which must include at least two industrial partners from different member states of the community. Besides industrial R&D projects, the program also includes basic research.

Mr. M. Moens, University of Edinburgh,
Centre for Cognitive Science, 2, Buccleuch Place, UK-EDINBURGH EH89LW
Tel: +44/31 667-1011 ext 6444

Following is a brief synopses on several ESPRIT projects as presented in Mr. Moens' article in DARPA's "Proceedings: Speech and Natural Language Workshop," February 1991.

- *Integration and Design of Speech Understanding Interfaces (SUNSTAR).* SUNSTAR's objective is to show benefits and enhancements that human computer interfaces offer when based on speech input/output. The project demonstrates this by achieving prototypes in two fields of speech application: professional, office-type environment and a public telephone network environment. These fields represent market sectors of rapidly growing importance. This project is application-driven in the sense that it concentrates on the integration of speech functions into demonstrator systems rather than on fundamental research issues of speech recognition and output.

  Dialogue design and associated ergonomic aspects are of high importance to the project, in order to gain wider acceptability for speech interfaces in real-world applications. Another key issue is the integration of speech technology with other input/output devices.

  Dr. Joachim Irion, AEG Olympia AG,
  Max-Stromeyerstr. 160,D-7750 Konstanz; Tel: +49/ 7531-818559

- *Adverse-Environment Recognition of Speech (ARS).* The objective of this project is to develop improved algorithms for speech recognition in the presence of noise, and to build a real-time demonstrator. The demonstrator is intended to verify algorithm performance and address the problem of speech-based man-machine dialogue as a system interface in practical applications.

  Two application environments have been chosen: vehicles and factories. The system will have a vocabulary of 100-5400 words, chosen by each national group of partners and tailored to specific application environments.

The real time demonstrator will be based on a general-purpose DSP chip attached to a personal computer or a stand-alone system. In the development system, the signal processor will be connected to a host which provides for development support of software algorithms and acts as a file server for the required databases. First, performance evaluations will be made in the laboratory using suitable databases collected in noisy environments, by measuring the resulting rate of correct recognition. Performance under field conditions will then be assessed from a prototype fitted in a vehicle and a laboratory system installed in a factory.

The project will interact with DRIVE research program projects dealing with vehicle applications, and with other European projects on speech recognition.

Progress and Results. After the project's first year, a multilingual database collected in noisy environments was made available between the partners and used for evaluation of their baseline systems. These systems were realized according to a common standard suitable for exchanging the software modules of the algorithms (studies of which are presently in progress) between partners. The hardware structure of the final real-time demonstrator has been defined.

Mr. G. Babini, CSELT, Via G. Reiss
Romoli 274, I-10148 TORINO; Tel: +39/11-2169391

- *Multi-Language Speech-to-Text and Text-to-Speech System (POLYGLOT)*. The goal of POLYGLOT is to demonstrate the feasibility of multi-language voice input/output for a number of commercially promising applications. The objective is to integrate phonetic, lexical, and syntactic knowledge common to text-to-speech and speech-to-text conversion, providing greater generality, lower cost, and easier extensions. This project is based on results of ESPRIT project 860. An existing isolated-word speech recognition system will be extended, under this project, to six other languages.

Progress and Results. Detailed speech database specifications have been completed for isolated word recognition, continuous speech recognition and text-to-speech. Full specifications of the POLYGLOT common hardware and software are also available. A tool for the acquisition and hand segmentation or labeling of speech, SAMBA, has been implemented on a PC. Also completed was a modular architecture for "time delay neural networks."

The first six months' work on text-to-speech was mainly preparatory and theoretical in nature. However, specifications for the following are now available: system architecture, automatic language identification, voice source and vocal tract model, analysis and development tools, prosody and intonation, and working environment for synthesis rule development.

Mr. V. Ing, C. Olivetti & C. SPA,
Corso Svizzera 185,I-10149 Torino; Tel: +39/11-748162

- *Speech Understanding and Dialogue (SUNDIAL).* SUNDIAL addresses the problem of speech-based cooperative dialogue as an interface for computer applications in the information services domain. The main technologies to be developed will include continuous speech recognition and understanding, and oral dialogue modeling and management.

Speech input will be sentences of naturally spoken utterances of telephone quality with a vocabulary of 1000-2000 words for each application. The grammar will be based on a subset of the four partners' languages (English, French, German and Italian). The project was started with speaker-independent recognition of sub-word units. The second phase will consider automatic on-line speaker adaptation with a view towards improving performance. The dialogue manager allows users to express themselves in a restricted natural language.

Prototypes will demonstrate technology for three main information service applications: intercity train timetables (German), flight information and reservations (English and French) and a hotel database (Italian). The spoken language phenomena to be addressed will be determined from analysis of both human-human dialogues as well as human-machine simulations. Each demonstration system will be evaluated through extensive user trials.

For all demonstrators, the project has to define a common general architecture, common formalisms for grammar representation across languages, and common semantic representations for dialogue management and message generation.

Progress and Results. The project started with a number of definition studies for the general architecture and studies of application scenarios. A common architecture has been defined, together with the interfaces between the major modules; this facilitates comparative evaluation and exchange between partners.

A small vocabulary of 50 words was developed for the telephone speaker-independent recognizer, which is suitable for a banking-by-phone application. Tests on the recognizer using the Recognizer Sensitivity Analysis (RSA) technique has shown 95.6% correct recognition (+/- 0.7% at the 95% confidence level) on the RSA 31-word vocabulary.

Preliminary results for the acoustic-phonetic decoding module shows that continuous density HMMs (CDHMMs) achieve 77.6% word accuracy on sentences compared to 68.5% for discrete density HMMs using 275 phonetic units for the Italian language and a near 1000-word vocabulary. These results are for speaker-independent recognition of telephone quality sentences, but do not take into account the effect of the linguistic processing module on sentence understanding performance.

Results for the English language using CDHMM show that phoneme recognition accuracy on the DARPA TIMIT database is comparable to that achieved by Kai-Fu Lee in the Carnegie Mellon SPHINX system.

A common dialogue manager architecture has been defined and work is in progress on its implementation. The first full working prototype oral dialogue system is on target for completion by July 1991.

Mr. Jeremy Peckham, LOGICA LTD, Betjeman House
104 Hills Road; Tel: +44/223-66343

- *Multilingual Speech Input/Output Assessment, Methodology and Standardisation (SAM).* The objective of SAM is to develop methodologies, tools and databases for the assessment of speech synthesis and recognition systems in application where multilingual performance is required from the same basic equipment. The consortium is necessarily broad, with participants from six European communities and two EFTA member states. The project is able to provide techniques for assessing speech synthesizers and recognizers for the eight languages of the participating countries.

The participation of such a large range of organizations ensures the final recommendations will be widely adopted. Furthermore, close ties have been established with related national projects in the participating countries, all of which are moving towards the use of SAM standards.

During the definition phase of this project, which was supported under ESPRIT project 1541, a first multilingual speech database was established on CD-ROM, and this continues to be widely used for the purposes of assessment, analysis and research.

Progress and Results. The activities in the present, main phase of the project focus around three major areas: (1) speech input assessment; (2) speech output assessment; and (3) enabling technologies.

Prof. A. Fourcin, University College London
Wolfson House, 4 Stephenson Way, UK-London NW1 2HE
Tel: +44/71-3871055

- *Robust Analytical Speech Recognition System (ROARS).* The goal of the ROARS project is to increase the robustness of an existing analytical speech recognition system (i.e., one using knowledge about syllables, phonemes and phonetic features), and to use it as part of a speech understanding system with connected words and dialogue capability. This system will be evaluated for a specific application in two European languages.

  The work starts from an existing system implemented for the French language. This system has been shown to operate in real time, is speaker-independent, and has had satisfactory results with continuously uttered connected words. The aim of the first phase of the project is to develop and implement the corresponding knowledge-bases for the Spanish language and to enhance the robustness of this system against:

  -- intra- and inter-speaker changes in articulation, by the improvement of knowledge used in the system, including the possibility of a progressive and slow automatic adaptation;

  -- various ambient noises, by analyzing the degradations induced on each feature, rules used in the phonemic recognition system and the changes in articulation (at the feature level) when the speaker is in different noise conditions; examining the problem of false alarms at the sentence detection level; and studying and testing improvements aiming to minimize these degradations.

  All these tasks will be run in parallel for both languages (French and Spanish). Two identical hardware prototypes will be built, (one for French and the other for Spanish), to study implementation and test improvements.

  The purpose of the second phase is the implementation of two demonstrations of speech understanding for air traffic control (one in French, one in Spanish) and the integration of voice input with other devices such as keyboards, trackerballs and screens. These demonstrations will require a vocabulary of 100 to 200 words, connected words, and multimedia dialogue.

  Mr. Pierre Alinat, Thomson-CSF/SINTRA-ASM
  Chemin des Travails, F-06801 CAGNES SUR MER; Tel: +33/92-023211

Outside of DARPA, City University of New York and IBM are involved in active research specifically concerned with automatic voice-to-TDD relays. Data from their research will start to become available early in 1992.

Advanced Technology Systems

The following paragraphs describe existing voice technologies implemented in hardware and software which can be purchased off-the-shelf. Because the list of voice recognition systems is extensive, we have selected a representative list of a few of the technologies to describe here in detail. Table 1 provides a fairly representative list of the present manufacturers and vendors of speech recognition systems in the United States and Europe.

Dragon Systems, Inc., 617/965-5200, is one of the leaders in real time voice recognition systems. Two of their voice recognition systems are:

- *IBM VoiceType*, a powerful, free-text, large vocabulary, discrete word, speech recognition system which provides hands-free access to PCs. This product is a result of a cooperative arrangement between Dragon Systems and IBM. It has a 7,000 word active vocabulary (5000 base words and 2000 slots for user-defined words); its 80,000 word back-up dictionary ensures accuracy and improves productivity. This system has built-in commands for controlling popular word-processing, spreadsheet, and database software programs. By using Dragon Systems' patented speaker adaptation feature, the system learns the user's voice. This system is a powerful cost-effective solution for people with physical disabilities that preclude manual typing. The list price is $3,185. The system requires an IBM Audio Capture and Playback Adapter (ACPA) Board, which is purchased separately. Other system requirements are: a minimum of 6Mb memory and a 30Mb fixed disk (60Mb fixed disk recommended); and, DOS 5.0.

- *Dragondictate™*, is a software program and peripheral card that plugs into a personal computer to transform the computer into a voice-driven typewriter with large vocabulary discrete word recognition capability. DragonDictate performs at an average rate of 30 to 40+ words per minute. It recognizes 30,000 words and also incorporates an 80,000 word dictionary. This programs learns a user's vocabulary and speaking style. System requirements are an IBM PC AT compatible computer with 8 megabytes RAM, hard disk with 10 megabytes, 5.25" high density disk drive or 3.5" disk drive, expansion slot for the speech recognition board and an Intel 80386 microprocessor running MS-DOS. The cost is $9,000.

On both of the above systems, the accuracy starts out low but builds as the user works with the system. Depending on the user, the recognition rate can be as high as 98 to 100 percent.

Table 1. Speech Recognition Manufacturers and Vendors

CANADA	
Applied AI Systems, Inc. 340 March Rd., Suite 500 Kanata, Ontario K2K 2EA	
EUROPE	
Deltatre Voice Connexion Intl. Via Nino Bixio, 8 Peschiera B. 20068 ITALY	Lernout & Hauspie Speech Products Koning Albert I Laan 64 1780 Wemmel BELGIUM
VECSYS Le Chene Rond Bievres 91570 FRANCE	
UNITED KINGDOM	
Aptech Collingwood House Meadowfield, Pointeland Newcastle Upon Tyne NE20 9SD	Macfarland Systems, Ltd. 34 Eden Street Kingston Surrey KT1 1ER
Telsis Limited Barnes Wallis Road Secensworth East, Fareham Hampshire PO15 5TT	
UNITED STATES	
British Technology Group USA Renaissance Business Park Gulph Mills, PA 19406	California State University, Northridge Office of Disabled Student Services 18111 Nordhoff St.-DVSS Northridge, CA 91330
Covox, Inc. 675 Conger Street Eugene, OR 97402	DAC Systems 16 Colony Street Shelton, CT 06484
Dragon Systems 320 Nevada St., Second Floor Newton, MA 02160	Electronic Telecommunications, Inc. 3620 Clearview Parkway Atlanta, GA 30340
Emerson and Stern Associates, Inc. 10150 Sorrento Valley Drive, Suite 210 San Diego, CA 92121	Gardient Technology, Inc. 95B Connecticut Drive Burlington, NJ 08106

Gralin Associates, Inc.
3605 Old Easton Road
Doylestown, PA 18901

Hearsay Inc.
307 76th St.
Brooklyn, NY 11209

InterVoice, Inc.
17811 Waterview Parkway
Dallas, TX 75252

ITT Aerospace/Communications Division
492 River Road
Nutley, NJ 07110

Kurzweil Applied Intelligence, Inc.
411 Waverly Oaks Road
Waltham, MA 02154

Linkon Corporation
226 East 54th Street #301
New York, NY 10022

Mimic/Perlex
1720 East Morris
Wichita, KS 67211

Periphonics
4000 Veterans Hwy
Bohemia, NY 11716

Scott Instruments
1111 Willow Springs Drive
Denton, TX 76205

Shure Brothers, Inc.
222 Hartrey Avenue
Evanston, IL 60202

Simpact Voice Products Group
1782 La Costa Meadows Drive
San Marcos, CA 92069

Speech Systems Inc.
18356 Oxnard St.
Tarzana, CA 91356

Street Electronics Corporation
6420 Via Real
Carpinteria, CA 9303

Summa Four, Inc.
25 Sundial Ave.
Manchester, NH 03103-7251

Syntellect, Inc.
15810 N. 28th Avenue
Phoenix, AZ 85023

Telephonics
789 Park Avenue
Huntington, NY 11743

V. Channel
713 Camina Escuela
San Jose, CA 95129

VoCal Telecommunications
77 West Las Tunas #202
Arcadia, CA 91007

Voice Control Products, Inc.
1140 Broadway, Suite 1402
New York, NY 10001

Voice Information Associates, Inc.
1775 Massachusetts Ave
P.O. Box 625
Lexington, MA 02173

Voice Processing Corporation
One Main Street
Cambridge, MA 02142

Voicesys
71 Mark Bradford Dr.
Holden, MA 01520

Votan
4487 Technology Drive
Fremont, CA 94538

Voice Connexion
17835 Skypark Circle, Suite C
Irvine, CA 92714

Soliloquy™ Language Recognition Software by Emerson & Stern Associates, Inc., 619/457-2526. This is a software-based speaker-independent, discrete word, speech recognition system. It can be used on a MAC IIci or over the telephone. Soliloquy accepts 20-200 words. Additional words can be added and the grammar can be modified. For short sentences (up to three words), the recognition rate is 85-95 percent; for longer sentences, the rate is approximately 70%. Emerson & Stern Associates is continually working on improving the recognition rate. Technical specifications include:

- Code: Written in C.
- Minimum processing power: equivalent to 1 M68040 chip at 25 MHz; 1 Mb RAM; actual amount needed depends on application.
- Input: digitized speech stream, at 11 kHz, 8-bit companded; can adapt to other rates and formats, e.g., 22 kHz, 8-bit linear or 8 kHz, 8-bit companded.
- Output: text transcription direct to screen and/or ASCII output over standard RS232 link or any other standard protocol.
- Response time: less than 2 seconds after end of speech.

Optional features and services include:

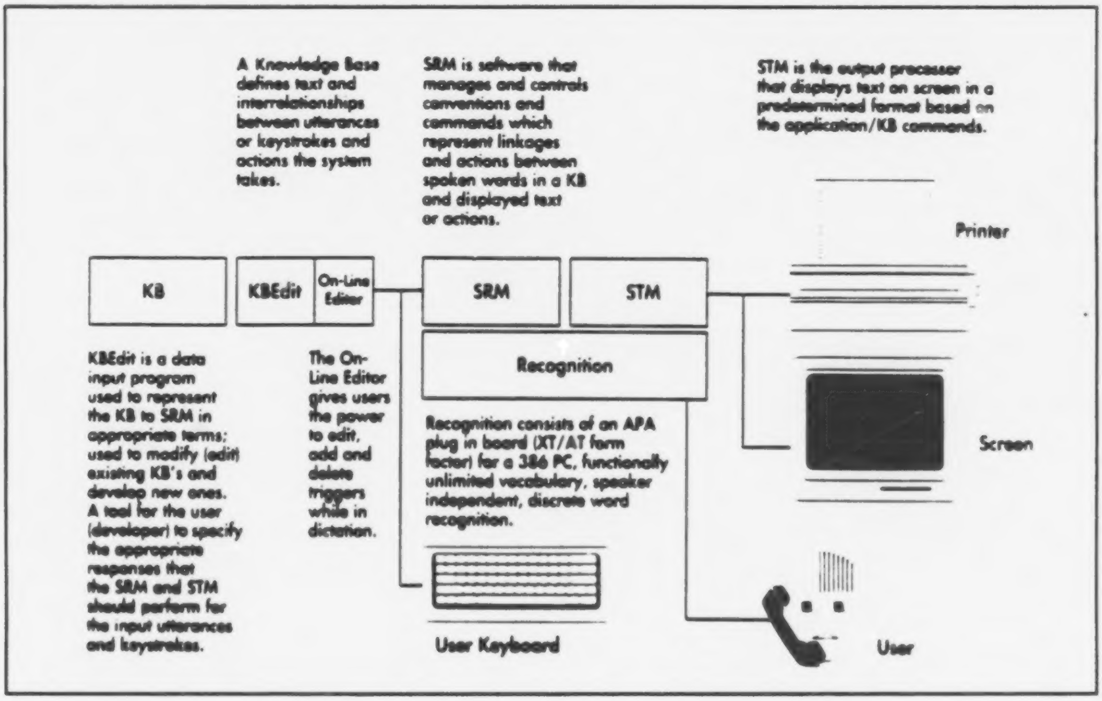
- Voice activation for Listen On/Off
- HyperCard compatible demonstration/integration modules
- Installation service
- Training service
- Vocabulary/Grammar Toolkit

It is worthy of note that Soliloquy evolved from Say and See™, a program for Macintosh computers that acts as a speech therapy tool for persons with speech disorders or hearing impairments. Say and See displays a profile of the speaker's head, but cut in half to show tongue movement as well as lip and jaw movement. The system displays this information solely based on "listening" to the person speaking. It is completely non-invasive.

VoiceReport by Kurzweil Applied Intelligence, Inc., 617/893-5151, is an open-ended development system for the creation of voice-generated documents. This system is speaker-independent and recognizes discrete words. It learns the voice of each user as well as the speaker's work usage pattern. Accuracy rate is 95 percent or better. As each word is spoken it is displayed on the screen. There is a functionally unlimited vocabulary size, and

VoiceReport can recognize homonyms. Documents can be edited by voice and printed simply by saying "print report." It uses both word-by-word dictation and trigger phrases; that is, a single spoken word or phrase can trigger a predefined sentence or paragraph.

System configuration consists of: 386-based host personal computer; 10MB high-speed 32-bit bus RAM; 40MB (or larger) hard disk; 1.2MB floppy; acoustics/phonetic analyzer card with digital signal processor (TMS320C25), 16 bit analog-to-digital resolution; and support for 14" color monitor; and it can be linked together on a Novell network. Figure 4 is a diagram of *VoiceReport* showing how the system works. The cost of the complete system, including the computer, board and software, is $18,900.



A Knowledge Base defines text and interrelationships between utterances or keystrokes and actions the system takes.

SRM is software that manages and controls conventions and commands which represent linkages and actions between spoken words in a KB and displayed text or actions.

STM is the output processor that displays text on screen in a predetermined format based on the application/KB commands.

KBEdit is a data input program used to represent the KB to SRM in appropriate terms; used to modify (edit) existing KB's and develop new ones. A tool for the user (developer) to specify the appropriate responses that the SRM and STM should perform for the input utterances and keystrokes.

The On-Line Editor gives users the power to edit, add and delete triggers while in dictation.

Recognition consists of an APA plug in board (XT/AT form factor) for a 386 PC, functionally unlimited vocabulary, speaker independent, discrete word recognition.

Figure 4. VoiceReport System Diagram

Voice Navigator II™ by Articulate Systems, Inc., 800/443-7077, allows control of the Macintosh by voice, using spoken commands to accomplish functions normally performed by using the keyboard or mouse. It is a discrete word system with 96 to 100 percent recognition rate. It works with any standard Macintosh computer application by providing language files containing basic voice commands for the popular applications. Also provided is the Language Maker™, a desk accessory, which allows creation of the user's own voice commands and language files for any application by simply pointing and clicking. The *Voice Navigator II* also allows the recording of voice or sound to be used for voice messaging over electronic mail networks, for voice and sound narration of HyperCard stacks, and for multimedia presentations.

Systems requirements include: Macintosh Plus, SE, II, or Portable with 2MB RAM; hard disk (preferably with at least 100KB space per user); microphone (preferably unidirectional or noise-cancelling); and System 6.0 or greater (compatible with MultiFinder). Cost for the system software is $795.00. The technical specifications are included in Figure 5.

Voice Connexion, 714/261-2366, offers three voice recognition systems. They are as follows:

- *PTVC-756*, a hand held IBM compatible computer with voice recognition and synthesis. This model is designed to handle most data collection, analysis and communication applications. The data acquisition device features voice recognition with 500 words per user and unlimited text-to-speech voice output for prompting and verification. The system recognizes discrete words and the recognition rate is 98%. A high contrast 16 line by 21 character display, a built-in serial port, and up to 1 megabyte of RAM memory are all housed in a high impact, environmentally-protected case. The cost for the system is $3195. Features and system requirements include:
    - Operating System--MS-DOS with full IBM XT compatibility.
    - RAM Memory--up to 1 MB for data, program files and application software.
    - Built-in RS-232/422 interface: data uploading/downloading to host; one or two-way acoustic communications.
    - Size--9.65" x 4.00" x 2.30"; Weight--34 ounces
    - Concurrent operation of voice recognition, bar code, and keyboard--operates transparently with application software.
    - Bar code--pencil wand or laser scanner.
    - Keyboard--50 key alphanumeric with standard IBM function set.
- *IntroVoice™* VI is a discrete word, voice recognition, and synthesis system for the IBM XT/AT/386 and PS/2 Model 25, 30, or compatible. It is a complete voice input/output system which provides voice recognition of 500 words with an accuracy of >98%, and unlimited text-to-speech synthesis. Introvoice VI listens to commands or data input and then responds by sending keystrokes

Processor/Memory
TMS 320C10 digital signal processor (DSP) with 16 bit data bus @ 14.3 MHz
8K words 100 ns static RAM memory
ROM memory for proprietary firmware

Sound Digitizers
8 bit μ law, companded CODEC and 8 bit linear A/D converter
Software selectable sampling rate up to 22.3 KHz

MACE Compression
On board, real-time support for standard Macintosh Audio Compression Expansion
22.3 KHz, 11.1 KHz, 7.4 KHz sampling rates.
1:1, 3:1, 6:1 compression ratios at all sampling rates

Audio Filters
Anti-aliasing low pass filters, software selectable
-3 dB @ 8 KHz and -3 dB @ 11 KHz

SCSI Interface
NCR 53C90 controller/driver
Max. transfer rate 256 Kb/sec (limited by Macintosh)
External SCSI termination and address selection switches
SCSI termination status indicated on front panel LED
25-pin SCSI input/output ports
Custom 25-pin to 25-pin SCSI cable (included)

Display Panel
4-segment LED indicator for audio input level
Bi-color LED indicator for power and SCSI termination status
Push button power switch mounted in display panel

Input Jacks
3.5 mm microphone jack accepts any standard microphone
2.5 mm microphone switch jack accepts any standard microphone switch
2.5 mm jack automatically supplies electret bias voltage for custom electret microphone
3.5 mm jack automatically operates as a microphone switch jack for custom electret microphone

Power Supply
External power supply module (included)
9 volts DC at 1 amp power supply, 110 volts, 60 Hz line voltage

Physical Dimensions
5.5"W x 6.3"D x 1"H
Weight: 7.7 oz
Orientation: Either horizontal on integral rubber pads or vertical with snap-in plastic feet

Figure 5. Technical Specifications for Voice Navigator II

to the computer and text to the on-board synthesizer for audio prompting and verification. Cost is $895.

Both voice input and output can be easily integrated with any standard application program, with no modification required to the existing software. Voice recognition and synthesis character strings are easily defined by the Voice Utility Program supplied with the system. Features include:

-- 500 isolated words/phrases per RAM resident vocabulary.

-- Up to 1000 keystrokes per spoken word.

-- Better than 98% accuracy for standard vocabularies. 99% for digit recognition.

-- Operates reliably in noisy environments in excess of 85 dB.

-- Operates concurrently with keyboard, mouse, tablet, or bar code reader.

-- Provides real-time prompting and verification without a visual display.

- *Home Automation Link (HAL)* is a system available for the Environmental Voice Control, using a personal computer. HAL can give independence to the physically handicapped. By voice commands they can operate the TV, turn lights on and off, and make phone calls. The IntroVoice VI board, mentioned above, is an integral part of the system and furnishes voice input/output with a voice recognition of 500 words per vocabulary and an unlimited text-to-speech synthesis. The HAL entry system includes:

-- Voice input/output
-- Telephone interface
-- Lights and appliance controller
-- Infrared remote controller
-- Software and manual.

The cost, including the IntroVoice VI board is $1495. Without the board, the cost is $700.

Series 7000™ Conversational Voice I/O System, by Verbex Voice Systems, 908/225-5225, is a continuous speech recognition system allowing users to capture data and perform transactions using a virtually unlimited vocabulary. The user trains on the system and must input each word that will be used. Recognition rate is 98-99%. The Series 7000 offers uses from high-powered applications, such as stock or commodities trading, to professional workstations, CAD/CAM engineering, and package handling. It has an active vocabulary of 2,100 words expandable to 10,000 words. Computer memory is the only factor limiting the size of the vocabulary. The system learns the user's voice. The Series 7000 design use the TMS 320C30 chip. As a stand-alone voice peripheral, the system can be connected to an existing computer system. Cost is $9600.

Voice Control Systems, 214/386-0300, offers several voice processing components. They are:

- *DVM-4™*, a multiple-channel, speaker-independent voice recognition board for voice processing systems.
- *TeleRec*, a board-level system providing a reliable and efficient means for remote information access and transaction processing over telephone lines.
- *Network Automation System*, a system that brings automation to operator-assisted telephone calls.
- *VoiceGateway*, a state-of-the-art Interactive Voice Response System.
- *CellDial™*, a voice recognition/voice response system designed for integration with cellular telephones.

Voice Control Systems generally licenses their products to third party companies.

Phonetic Engine Speech Recognition System, by Speech Systems, Inc. (SSI), 818/881-0885, is a stand-alone unit connected to an RS232 port on a workstation. SSI's *Phonetic Decoder™* software runs on the workstation and processes the output of the Phonetic Engine, converting it to text. An application program, running on the workstation, interacts with the Phonetic Decoder through the Phonetic Decoder Interface, a standardized set of subroutine calls; thus, the speech recognition capability is fully integrated into the application software. The Phonetic Engine also has built-in voice record and playback which is useful for prompts or voice storage. It can record and recognize the same speech simultaneously.

The key to this advanced commercial system is proprietary technology that represents speech efficiently. The speech is processed by a generic male or female "speaker model" into a phonetic representation (a sequence of "phonetic codes" designed to capture the underlying phonemes, or parts thereof, in a spoken sentence). Since phonemes are the basic speech sounds, this processing is designed to retain only the information relevant to recognizing the words spoken. The Phonetic Decoder software then uses a dictionary and grammar to efficiently translate this phonetic representation into text.

The speaker models use decision trees, an efficient form of simulated neural network, to create the phonetic codes. The codes are not explicit decisions on the phonemes, but instead are interpreted statistically by the Phonetic Decoder; the interpreting algorithm can be viewed as using a form of Hidden Markov Model which is designed to deal with the time units represented by the phonetic codes. SSI will provide technical papers which discuss its underlying technology in more detail.

10.0 COST CONSIDERATIONS OF VOICE RECOGNITION SYSTEMS FOR PERSONAL AND MEDIA ACCESS

Table 2 shows the prices of the current advanced technology voice recognition systems. Prices have been stable over the past year. This price stability is primarily due to the lack of competition in the various product offerings (i.e., each product is targeted to a different application or target consumer market). As competition increases, the cost of voice recognition systems is expected to decrease as with other computer-related equipment. In particular, as the second and third generation products begin to appear, the cost of the technology will be driven down by market forces and microelectronics implementations of voice recognition hardware.

Table 2. Cost of Voice Recognition Systems

SYSTEM	MANUFACTURER	COST
IBM Voice Type	Dragon Systems, Inc.	$3,185
Dragondictate™	Dragon Systems, Inc.	$9,000
Soliloquy	Emerson & Stern Associates, Inc.	$1,000*
VoiceReport	Kurzweil AI, Inc.	$18,900**
Voice Navigator II™	Articulate Systems, Inc.	$795
PTVC-756	Voice Connexion	$3,195 (512K) $3,695 (1MEG)
IntroVoice™ VI	Voice Connexion	$895
Home Automation Link (HAL)	Voice Connexion	$1495 w/Intro-Voice VI board+
Series 7000™ Conversational Voice I/O System	Verbex Voice System	$9,600

*Sold only in large quantities of 250 or more.
**Includes a complete system (hardware and software)
+$700 without the IntroVoice VI board

Incorporating voice recognition capabilities into devices such as TDD phone relay, closed caption or interpreter services will require a substantial investment that may not be practical for manufacturers of voice recognition systems to invest in without government assistance or sponsorship for the initial research and development phases. The reason for this is that the handicapped market is small which makes it more difficult to recover development costs within a production run without passing the full cost on to the consumer. The first applications will therefore be systems adapted from mass market devices such as transcription systems for doctors. With a systematic development approach to developing interfaces and databases for applications for persons with hearing impairments, the Department of Education can help reduce the cost of voice recognitions to meet the needs of persons with hearing impairments.

11.0 COST BENEFITS OF EARLY SPONSORSHIP OF VOICE RECOGNITION REQUIREMENTS FOR PERSONS WITH HEARING IMPAIRMENTS

The cost benefits associated with early Department of Education sponsored research and development for application to persons with hearing impairments is that the costs associated with this development will not have to be passed on to the user in the final product. The research and development areaS for this targeted research should be vocabulary database development and structuring, interface requirement definition, human factors determination, and marketing and dissemination of information on potential uses. This will simplify integrating the needs of persons with hearing impairments into the voice recognition systems and reduce the development cost to manufacturers.

More accurate cost projections on advanced technologies that are in the laboratory will have to await systems transition to the consumer electronics market. However, many of the systems are being implemented on IBM-PC 486 level machines with special digital signal processor boards. This would indicate that the hardware cost will decrease based on mass market demand. The real cost will be in the development of software and special applications to meet the needs of persons with hearing impairments. Small numbers translate to high costs for the software when amortized over only a few thousand items. Therefore, it is anticipated that voice recognition systems to meet the specific needs of persons with hearing impairments could range from 2 to 10 times the cost for the mainstream consumer market. Early research and development efforts by the Department of Education could significantly reduce these estimated multipliers by mitigating the risk to potential manufacturers.

12.0 PRESENT GOVERNMENT INVOLVEMENT IN ADVANCED VOICE RECOGNITION TECHNOLOGY

The U.S. Government involvement in voice recognition systems has been broad and includes National Security, Transportation, Commerce and Educational applications. To date the most significant unclassified advanced technology effort is being conducted by DARPA's Information Science and Technology Office. DARPA has fostered research and development in speech and natural language systems for over 20 years. The DARPA work has generated interest throughout the Government and the civilian community. In February 1989, DARPA began an annual review process via a Speech and Natural Language Workshop. In recent years, the DARPA speech and natural language workshop has evolved from a small informal discussion of current progress and plans to a much larger and more formal meeting which has become a primary forum for the exchange of major research results. The success of the workshops and the underlying DARPA-sponsored program of research in spoken language has been confirmed by the recent surge of interest in spoken-language systems outside the DARPA community. Section 9.0 of this scenario presented a description of the voice recognition systems being developed under this program. The program encompasses not only voice recognition but the entire range of spoken language, including sentence structure, formatting and usage, etc. DARPA's goals are to:

- establish some common reference points by assessing the current state of the art in both speech recognition and natural language processing;
- cross educate researchers in the relevant disciplines farther from their area of expertise:
- highlight areas of common interest, namely prosodics, spoken language systems, and development of shared resources;
- present current research results in both speech recognition and natural language processing.

In addition, the DARPA workshops provide a forum for a review of the work being done in Europe, which was presented here to provide an understanding of the efforts going on worldwide in speech recognition and natural language processing.

Most of the systems are in research and development and point to the three to five year applications area. Now is the time the technology can be influenced to add special features to meet the needs of persons with hearing impairments.

13.0 ADVANCED VOICE RECOGNITION TECHNOLOGY TIMELINE

Dragon Systems and other manufacturers have begun selling the first generation o of speaker trained speaker independent systems in 1991. In addition, advanced voice recognition technology is moving from the laboratory to the market place. Within the next one to two years, several user-independent continuous voice recognition systems are expected to be marketed based on the research sponsored by DARPA and private companies, such as, the America Telephone and Telegraph Corporation. With the development of microphone array systems and noise cancellation systems to reduce background noise, voice recognition systems can be applied to a number of interpreter services for persons with hearing impairments over the next three to five years. Hearing-impaired persons with residual hearing may also benefit from access to the audio output from the noise reduction systems, with or without the voice recognition output.

One service that could be prototyped within one to two years is live closed caption via voice recognition systems, for local news programs. This could be accomplished using one of the DARPA-sponsored voice recognition systems. For example, Stanford Research Institute's (SRI) voice recognition system, described in the advanced technology section could be used to caption a local news program in California on a trial basis without removing the equipment from the laboratory. Telephone lines could be used to send the captions to the studio closed captioning equipment.

Voice recognition technology is expected to mature over the next 5 years to the point where it will provide transcription, computer control, and interpreter services for persons with hearing impairments. What is needed is a comprehensive program to apply the technology to meet specific needs of persons with hearing impairments. This will require that databases be established, training programs formulated and specific goals set to allow the technology to be adapted for use by persons with hearing impairments.

14.0 PROPOSED ROAD MAP FOR INCLUSION OF VOICE RECOGNITION CAPABILITY IN ADVANCED TECHNOLOGY SYSTEMS

The Department of Education should begin the process of developing voice recognition technology for use by persons with hearing impairments by participation in the DARPA sponsored Speech and Natural Language Workshops beginning in the winter of 1991. This should be followed by the appointment of a Voice Recognition Advisory Committee round table to recommend specific goals for developing the technology into devices for persons with hearing impairments for the Department of Education. An extensive applications program should then be initiated to apply the technology to the specific applications defined by the Voice Recognition Advisory Committee, such as closed captioning, TDD relay services, and interpreter services in classrooms and meetings. It is expected that a five year, one million dollar per year effort will be required to develop the technology into prototype products for use by persons with hearing impairments. Innovative Grants, Small Business Innovative Research Grants, and specific applications oriented programs should be initiated to achieve the goals defined by the Department of Education.

The payoff at the end of five years is to empower the hearing impaired with systems that allow them equal access to television and telecommunications medias as well as access to personal communications services.

15.0 POTENTIAL PROGRAM SCHEDULE

Figure 6 is a Proposed Schedule for advanced voice recognition technology development to meet the needs of persons with hearing impairments. To ensure that the needs of persons with disabilities are considered, the Department of Education should participate in DARPA's Speech and Natural Language workshops. In addition, the Department of Education should form an Advanced Voice Recognition Development Committee, working with Government agencies, universities, voice recognition companies and the telecommunications industry to incorporate the requirements of voice recognition capabilities into new systems for TDD relay services, closed caption systems and interpreter services as early as possible.

In particular, the Department of Education needs to continue to identify specific needs and applications for voice recognition systems to meet the needs of persons with hearing impairments. A comprehensive program would include the following:

- description of the target audience;
- voice databases for each application, for testing new systems and applications;
- investigation of the noise components unique to each application;
- input techniques for voice applications (microphone array techniques);
- system interfaces (i.e., TDD, ASCII modems, computer systems etc.).

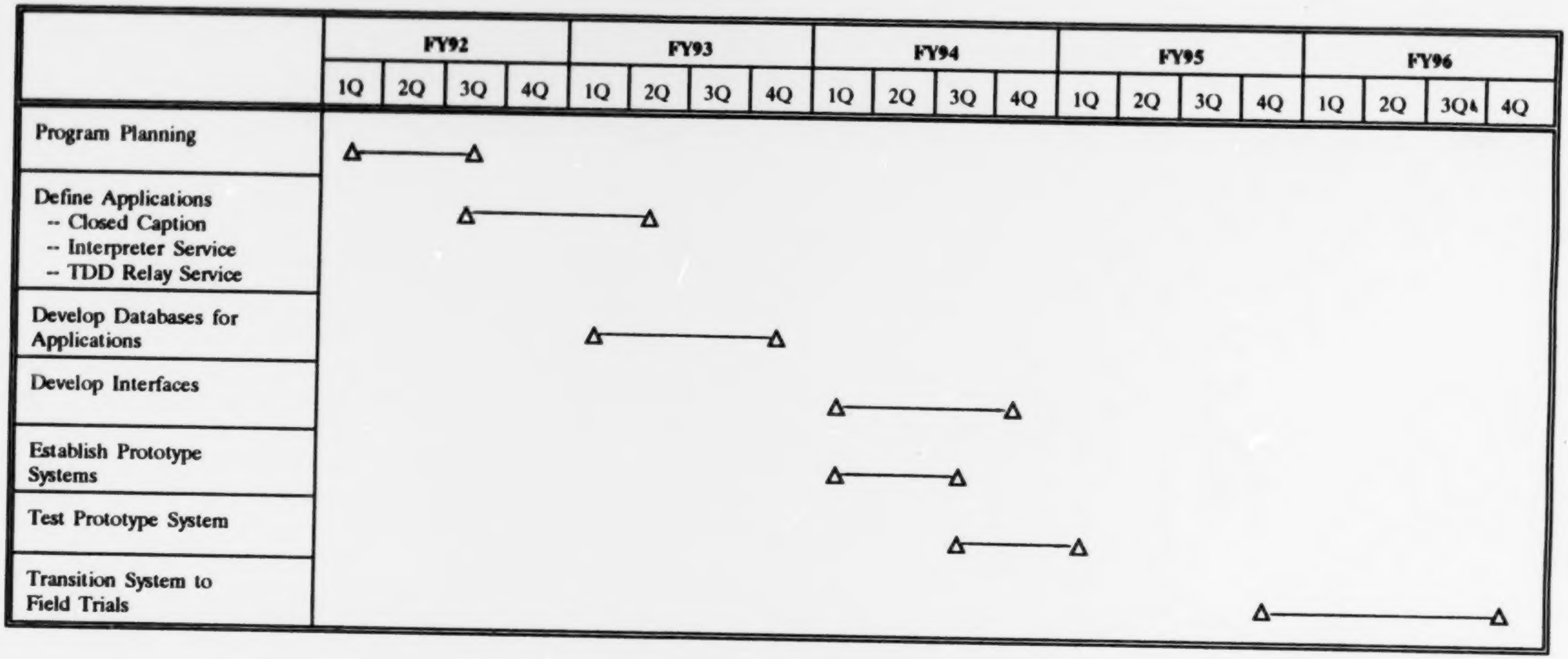


Figure 6. Timeline Schedule

VIDEO TELECONFERENCING/DATA COMPRESSION FOR PERSONS WITH HEARING IMPAIRMENTS

MARCH 1992

Prepared by

Daniel E. Hinton, Sr., Principal Investigator
and
Charles Connolly and Paolo Basso-Luca

SCIENCE APPLICATIONS INTERNATIONAL CORPORATION
3701 N. Fairfax Drive, Suite 1001
Arlington, VA 22203
(703) 351-7755

1.0 SCENARIO

Video Teleconferencing/Data Compression for Persons with Hearing Impairments.

2.0 CATEGORY OF IMPAIRMENTS

Persons with hearing impairments.

3.0 TARGET AUDIENCE

Consumers with Hearing Impairments. The consumer with hearing impairments will benefit from enhanced telecommunication system access through video teleconferencing, made possible by data compression. This scenario provides a means for persons with hearing impairments to potentially communicate over a telephone line or computer network using sign language. In particular, it provides a better understanding of the video compression technology available in electronic media over the next three to five years and the potential problems that could arise in telecommunication system access.

Policy makers, including national representatives, Government department heads, and special interest organizations. Policy makers can use this scenario to better understand the issues related to telecommunications system access for persons with hearing impairments. They may also use it to understand how advanced video compression techniques can make it possible to use existing facilities to easily create a new electronic phone service which can be used by both persons with and without hearing impairments.

Researchers and Developers. This group will benefit through a better understanding of the needs of persons with hearing impairments and specifically their telecommunications system access requirements. This understanding of telecommunication system access will assist researchers and developers in designing access functions in their future products to meet the needs of persons with hearing impairments.

Manufacturers. Manufacturers will benefit through a better understanding of the potential market size and the existing need for telecommunications access which can be met by video teleconferencing through data compression.

4.0 THE TECHNOLOGY

Many of the applications of digital video hinge on the use of image data compression, which means representing images in a more compact way to reduce the bandwidth required to send the images. Compression algorithms fall into two principle categories: information lossless and information lossy. Lossless compression means that in the absence of communication, the original image can be reconstructed *exactly* at the receiver. Information lossy compression, on the other hand, means that some error is introduced by the compression process itself. The objective of image compression algorithm development is to minimize the visual impact of these errors.

This section discusses methods for image data compression. All of the algorithms to be discussed here assume that the image data has been sampled and quantized at acceptable resolutions in both space and intensity. Initial intensity quantization of the sensor/scanner data is necessarily lossy but this is a property of the sensor, not the compression process. Lossless compression means that the *quantized* data is encoded and decoded with no loss of information due to the compression algorithm.

A taxonomy of popular compression schemes is shown in Figure 1. Those marked with a "*" are briefly described in the following sections. There are many compression algorithms not listed in Figure 1 and there are also many variations of the algorithms that are listed that we have chosen to ignore. The algorithms described below, however, form the foundation for most system needs.

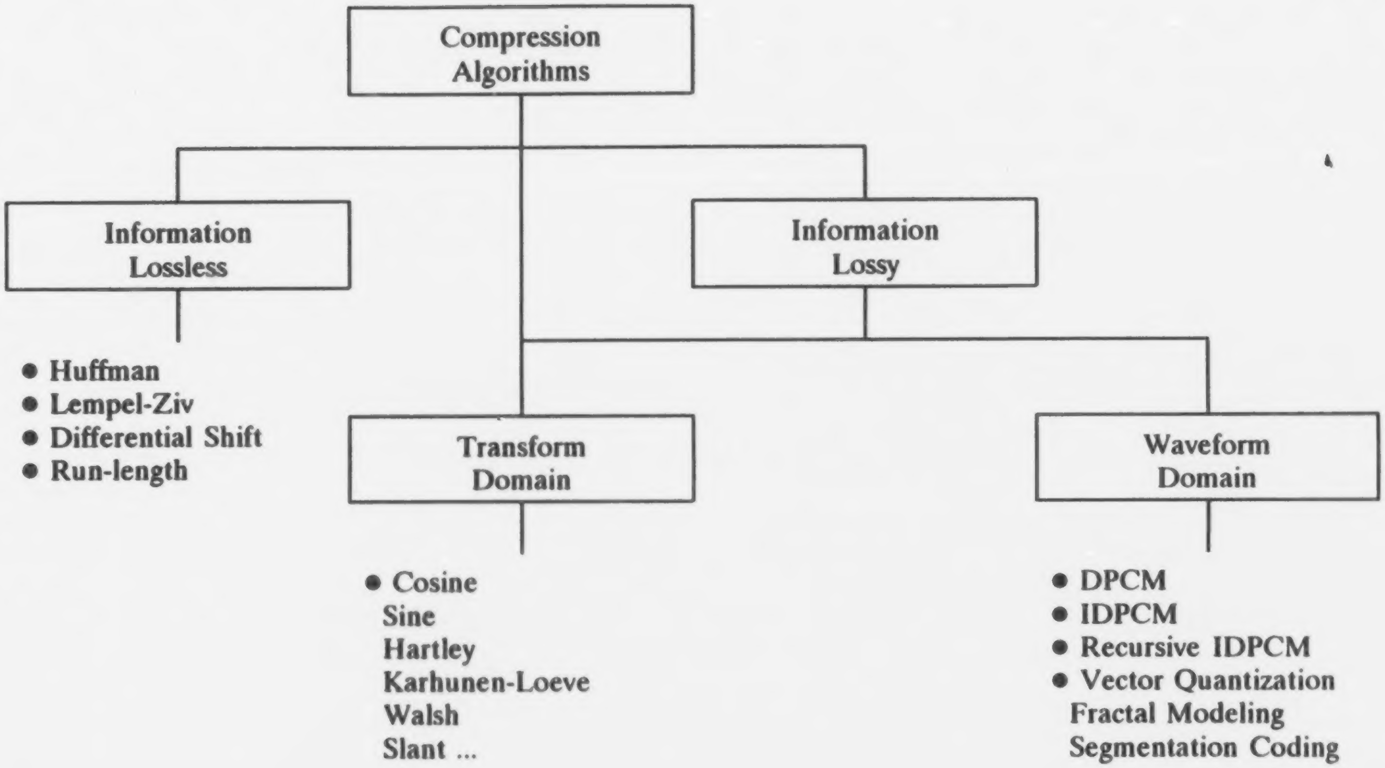
4.1 Lossless Coding

In some applications it is desired to compress the data subject to the constraint that the encoding process be reversible (i.e., no information is thrown away). We consider three cases of lossless encoding. The first two, Huffman and Lempel/Ziv can be applied to virtually any kind of data, whereas the third, Differential Shift encoding, is specifically for signals such as speech and image data having local spatial or temporal correlation. We further note that all lossy algorithms produce a sequence of channel symbols that can be further compressed using lossless techniques. That is why lossless techniques are relevant to image transmission over transmission channels as narrow as phone lines, despite the modest amount of compression they provide.

4.1.1 Huffman Coding

Huffman coding[1] is a methodology for assigning variable-length binary codes to fixed-length blocks of data. The length of the symbol assigned to a particular block of data is roughly inversely proportional to the statistical relative frequency of occurrence of that block. For example, if Huffman coding is used to encode single characters of English text, then short codes will be assigned to the frequently occurring letters "e" and "s" and long codes to the less frequent "z" and "q." Codes can be assigned, instead, to a sequence of K characters in combination. In this case there must be a channel symbol for each combination of K characters that can actually occur.

The drawback of the Huffman code, and others like it, is that they require the set of probabilities to be estimated from typical data. They could be determined directly from the data to be compressed but this requires an extra pass through the entire data set to compute a histogram of the number of occurrences of possible graphical elements. The alternative is to apply a fixed Huffman code to all incoming data.


Compression Algorithms
Information Lossless
Information Lossy
• Huffman
• Lempel-Ziv
• Differential Shift
• Run-length
Transform Domain
Waveform Domain
• Cosine
Sine
Hartley
Karhunen-Loeve
Walsh
Slant ...
• DPCM
• IDPCM
• Recursive IDPCM
• Vector Quantization
Fractal Modeling
Segmentation Coding

NOTE: Those with *'s are reviewed in this document.

Figure 1. Taxonomy of Compression Algorithms

4.1.2 Lempel/Ziv Coding

Lempel/Ziv (LZ) coding[2] is a lossless compression technique that is the basis for the UNIX utility *compress* and is one of the methods commonly used in file archival programs. Whereas the Huffman scheme generates a variable length code for fixed length sequences of characters, the LZ code creates fixed length codes (codewords) for variable length sequences of characters. An advantage of LZ is that the code is generated "on the fly;" it does not require a separate pass to collect statistics. Secondly, the fixed-length code makes it easier to design and implement subsequent channel error correction techniques.

A variation of the LZ algorithm, called Lempel-Ziv-Welch (LZW)[3] is described here. LZW incrementally builds a string table that is used to translate incoming strings to outgoing codewords. The table has the property that if the string wK, composed of some string w and a single character K is in the string table, then the prefix w is in the string table. The encoding algorithm is as follows:

```
(0)   Initialize the table to contain all possible single character strings. This is the
      underlying alphabet. 2 = first input character.

(1)   Read the next character K.

            If wK is in the string table:          w = wK, repeat (1).

            else wK is not in string table:        output codes (w)
                                                   add wK to string table
                                                   w = K; repeat (1).
```

An example of LZW encoding is shown in Figure 2. The LZW algorithm is very fast when implemented in hardware (UNISYS has proprietary hardware and software designs). The decoding algorithm is similar to the encoding algorithm but faster since there is no need to search the string table; the codeword can be used as an index to the corresponding character string. In order to avoid ambiguities during decoding, however, the logic required for decoding is slightly more complicated than for encoding. Both the encoder and decoder begin with a known set of codes for single characters. Compression ratios obtainable with LZW depend on the source data. Ratios of .4 to .6 have been achieved with image data.

4.1.3 Run-Length Coding

One of the simplest of the lossless algorithms is run-length coding.[4] In this method, adjacent picture elements (pixels) are compared in gray level, and only changes that exceed a threshold are encoded. In order to be lossless, the threshold is zero. Each scan line in a picture is coded as a sequence of ordered pairs of numbers, each pair containing a gray level and a run length. Since some run lengths occur more frequently than others, enhanced compression is achieved by Huffman coding the run-lengths.

String Table	a	b	c	ab	ba	abc	cb	bab	babc
Code	1	2	3	4	5	6	7	8	9
Time Added	0	0	0	2	3	5	6	8	11

position (time)	input symbol	output code	new string added
1	a	1	
2	b	2	ab → 4
3	a		ba → 5
4	b	4	
5	c	3	abc → 6
6	b		cb → 7
7	a	5	
8	b		bab → 8
9	a		
10	b	8	
11	c		
12	b	7	babc → 9

NOTE: LZ creates fixed codes for variable length sequences of characters.

Figure 2. Construction of a Lempel/Ziv Code

Run-length coding is an efficient method of compressing binary image data, such as drawings or text. For multilevel image data, however, the run-lengths tend to be very short. It is possible, under these circumstances, for compression to actually increase the size of a file. It would seem that map data would be a natural candidate for run-length coding because map data is symbolic and contains large regions of "uniform" color to represent different types of terrain. However, the presence of half-tone screens, along with noise in the scanning process, reduces the efficiency of run-length coding.

Run-length coding, like Lempel/Ziv coding, can be used with another compression technique like quantization to obtain further lossless compression.

4.1.4 Differential Shift Coding

Differential Shift Coding[5] takes advantage of the fact that adjacent pixels in an image (or samples in a speech signal) tend to be highly correlated. Due to this correlation, the dynamic range of the differences between the gray levels of adjacent pixels is considerably less than the dynamic range of the original image and can therefore be represented by fewer bits. The procedure is best described by example.

Suppose we construct a 4-bit code consisting of 14 codewords, c_2 though c_{15}, representing the differences -7,-6,-5,...,-1,0,1,...,6 respectively. As long as differences remain within this range, we can transmit at four bits per pixel. To accommodate larger differences, we use c_1 and c_{16} to shift the range of differences. c_1 is used to indicate that the pixel difference is less than -7 and is interpreted as "subtract 14." c_{16} is similarly interpreted as "add 14." Thus a difference of 13, for example, would be represented by the pair $c_{16}c_8$ (c_8 is the code for -1). Larger shifts can be handled by repeated shift codes, e.g., $c_1c_1c_1c_1c_1c_6 = -73$. This particular code requires 24 bits but it rarely occurs for typical data since most of the differences are within the range -7,...,6. If the statistics of the symbols are known, then they can be Huffman encoded. This is known as a Huffman shift code.

4.1.5 Drawbacks of Lossless Techniques

All lossless techniques produce a variable length code, i.e., one cannot predict the exact number of bits required to code a particular data set. If the coded data is to be sent across a communication channel, it is necessary, therefore, to provide buffers and a mechanism for dropping data (say at the end of a scan) should the buffers overflow.

Differential Shift and LZW coding are also sensitive to channel errors because the characters or values reconstructed by the decoder depend on previously decoded values. An error puts the decoder in an incorrect state which can persist until the decoder is reinitialized. Differential Shift coding can be reinitialized at the start of each image scan line. In this case, a channel error produces a streak or gap in the scan line following the location where the error occurs but it does not propagate to the next line. There are techniques for filling these gaps using interpolation from adjacent lines but the gaps must first be detected. These operations add considerable complexity to the system.

A channel error occurring in an LZW system can cause drastic output errors because the decoder is a finite state machine whose state depends on previously decoded symbols. The algorithm also loses its compression efficiency if it has to be restarted too often. Blocks of 10^3 to 10^4 characters are recommended for efficient use. If the channel bit error rate is above 10^{-7}, this method is not recommended. Since bit-error rates for CD-ROMs are typically better than 10^{-12}, LZW is definitely applicable.

The principal drawback of lossless algorithms is that they are unable to achieve high compression rates. Experiments on image data rarely yield compression ratios greater than 2.5:1.

4.2 Information Lossy Algorithms

When transmitting or storing data that is to be interpreted by the human visual system, minor differences between the source image and reconstructed image are often insignificant. The goal in lossy image compression is to distribute the inevitable error in such a way as to minimize visually perceptible artifacts. The error is produced by requantization of the data, but only after the data has been manipulated to decrease the dynamic range and increase the statistical independence of the samples.

Lossy compression algorithms can be divided into two categories, spatial domain and transform domain, depending on where requantization of the data occurs. In the next section, we briefly review the topic of scalar quantization. Discrete Cosine Transform (DCT) coding, a popular representative of the transform domain class, is then described. In the spatial domain category, Differential Pulse Code Modulation (DPCM), Interpolative DPCM (IDPCM), and Vector Quantization (VQ) will be discussed. IDPCM is the algorithm adopted for the National Image Transmission Format (NITF).[5] NITF was designed primarily for secondary dissemination of image data and accompanying text.

4.2.1 Scalar Quantization

Scalar quantization is the process of representing a numerical quantity by a finite number of bits. If b bits are used to represent a quantity V, then V can take on one of 2^b different values. These values, called the quantization *levels*, are usually selected carefully to minimize a distortion criterion. The quantization error is introduced by approximating each sample by one of the 2^b levels.

All of the information lossy algorithms use some form of quantization to reduce the number of bits used to represent the data. The trick is to apply quantization to a transformation of the data that reduces both the sample to sample correlation and the dynamic range of the variables to be quantized. All of the algorithms described below, with the exception of Vector Quantization (VQ), use scalar quantization. VQ is simply a generalization of the quantization process to multi-dimensional variables (vectors).

4.2.2 Transform Coding

Transform domain algorithms are motivated by the desire to perform scalar quantization of uncorrelated data, i.e., data for which the spatial redundancy has been removed. The optimal algorithm for obtaining truly uncorrelated samples is the Karhunen-Loeve (KL) transform;[6] which is based on the covariance matrix for each data set to be compressed. Since KL is computationally expensive, it is often approximated in practice by the Discrete Cosine Transform (DCT) for which a fast transform algorithm exists. In fact, one can use the Fast Fourier Transform (FFT) to compute the DCT (but it is not the most efficient method).

4.2.3 Differential Pulse Code Modulation

Differential Pulse Code Modulation (DPCM) is a spatial domain compression system in which the difference between actual and predicted image values are scalar quantized and transmitted. The basic structure of a DPCM system is shown in Figure 3. The value $\hat{f}(x)$ is a prediction of f(x) based on previously quantized prediction errors. The injected quantization error prevents exact reconstruction of the source data.

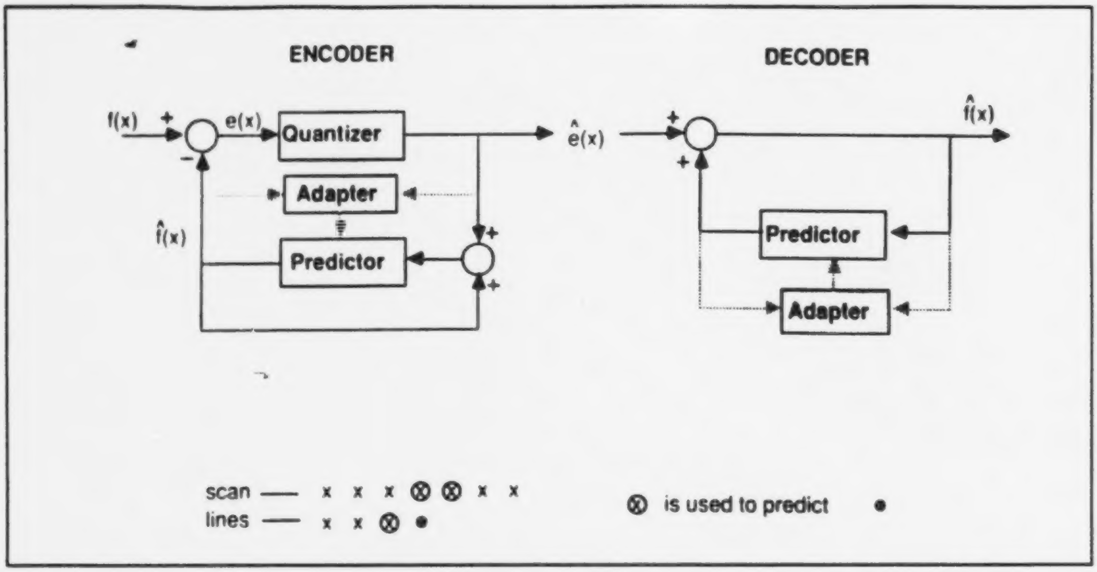


Figure 3. Differential Pulse Code Modulation Block Diagram

There are numerous variations on the DPCM theme. Recognizing that image statistics are not stationary, most of these variations address methodologies for making DPCM adaptive. Adaptivity always adds computational complexity. Fortunately there are some attractive alternatives, including LMS Adaptive DPCM and IDPCM.

4.2.4 LMS Adaptive DPCM

A simple and effective means of allowing the DPCM system to adjust to local statistics is to allow the predictor coefficients {h's} to follow a least mean squares (LMS) updating rule.[7] With each quantized residual $\hat{e}_k$, the predictor coefficients are updated according to

$$h_k^{(n+1)} = h_k^{(n)} + \alpha \hat{e}_k \hat{f}_k^{(n)} \qquad k=1,\ldots,K$$

where the superscript indexes spatial location and k indexes through the samples used in the predictor.

4.2.5 Interpolated DPCM

The DPCM algorithms previously described predict the value of the next sample based on samples previously encountered during scanning of the data. These are known as "causal" algorithms because they use only past events to predict current or future events. The ordering of the image data into a temporal sequence is only a convenience--it is not necessary. The Interpolated DPCM (IDPCM) algorithm is a DPCM algorithm which uses several samples in the spatial neighborhood of a sample to predict the value of that sample. The prediction algorithm uses simple linear and bilinear interpolation.

4.2.6 Vector Quantization

Vector quantization (VQ) is a compression technique in which blocks of data are quantized jointly. The blocks of data, consisting of N scalar samples, are treated as vectors in an N-dimensional space. The quantization levels in an N-dimensional vector quantizer are also N-dimensional vectors which are usually chosen to minimize a selected error measure over a training set. In the jargon of VQ, each quantization *level* (actually a vector) is referred to as a codeword and the set of codewords is called the codebook.

4.3 Algorithm Design Issues

Published results on the application of lossless algorithms to image data show that the compression ratios average about 2.2:1. Thus compression ratios of 500:1 or greater required to transmit sign language over a phone line, require incorporating a lossy algorithm. Transmitting sign language over computer networks, however, requires far less compression, sometimes requiring only reduced resolution; compression helps though. The actual compression rate achieved by a lossless algorithm cannot be predicted--it is data dependent. This is also true for any adaptive lossy algorithm which uses Bit Assignment Matrices (BAMs) or other variable quantization schemes, producing spatially variant block compression rates. A computer network can often handle this variation well, but variable data rates present a challenge when used over a phone line.

4.4 Image Fidelity

There are two types of fidelity criteria; objective and subjective. Algorithm design is usually based on objective criteria (such as the minimization of mean squared encoding error) since they can be used to form cost functionals to be minimized during design of the encoder. Subjective image quality studies have repeatedly shown, however, that subjective and objective criteria are only partially correlated. In addition, there is no general agreement regarding the suitability of subjective criteria; they are almost always adjusted to meet the needs of the imagery "consumers." Although the functional utility of the data should be tied to the subjective evaluation criteria, it is not easy to do so. When presented with paired comparisons of original and reconstructed data, it may be possible to see changes in color or texture which allow an individual to discriminate between the original and reconstructed data, but there may, in fact, be no degradation in functional utility. The difficulty of these issues has prevented the establishment of a standard way to judge effective image quality. Some of the factors that determine image quality include:

- Color fidelity
- Clarity of contour lines, text, and boundaries
- Pattern and shading reproduction

- Elimination or degradation of important details
- Introduction of invalid texture or granularity.

5.0 STATEMENT OF THE PROBLEM

5.1 Introduction

Approximately 2 million Americans have hearing impairments severe enough to make speech unintelligible with a hearing aid. Of these, "about 200,000 were born deaf or became deaf before they learned a spoken language, about 410,000 became deaf before the age of 19 years, and most of the remainder became deaf in later life as an all-too-common concomitant of aging."

"American Sign Language (ASL, Ameslan)--the sign language now in common use in the U.S.A. and Canada--[enables] deaf persons to communicate with each other on everyday, nontechnical subjects at about the same speed as hearing persons communicate in ordinary speech. ASL is in no sense a copy of English--it has its own distinctive grammar and modes of expression."

"Words such as proper names or technical terms for which no sign yet exists in ASL can be expressed by means of finger spelling--a letter-by-letter rendition of the word by means of a stylized set of finger positions. The rate of finger spelling normally is several letters per second with skilled users approaching a rate of 10 letters per second. Nevertheless, finger spelling is much slower than either spoken language or ASL."

"Sign languages enable the deaf to communicate...with great facility, in contrast to the difficulty with which the deaf communicate with the hearing community by means of reading lips and facial expressions, and by means of written messages. [Because] it can be easily learned and greatly speeds communication, ASL is known to the majority of congenitally deaf adults regardless of their educational background."[8]

5.2 Sending Sign Language Over A Standard Telephone Line

Telephone connections are not standard because older telephone wires and cables have much more noise than the latest fiber optics and microwave channels. Thus, *standard* is used here to mean *available anywhere in the U.S. at no extra charge*. A *standard phone line* is used here to refer to a *single voice channel*, used to carry one conversation (as opposed to multiple voice channels, which may be carried on a single cable or optical fiber).

Two devices that are providing telecommunication for the deaf are telephone devices for the deaf (TDDs) and the video telephone. TDDs permit a sender to type messages to a receiver who sees the characters displayed on a screen or produced on another TDD. Although TDDs are useful for communication between deaf and hearing people, they have a practical disadvantage in that communication is slow and effortful when compared with voice or ASL communication. Even finger spelling can reach 10 letters per second which is equivalent to a typing speed of about 120 words per minute (wpm). Baudot TDDs, the most common type, impose a hard limit just below 80 wpm. Eighty wpm rarely imposes any practical limitation, but Baudot TDDs only permit communication in one direction at a time, severely impairing the normal flow of conversation. ASCII TDDs, which are essentially computer modems with a keyboard and screen, permit two-way simultaneous communication limited only by typing speed, but typing speed and conversational human factors issues associated with communication by typing are inherent to all TDDs.

The video telephone is far more attractive than the TDD to many deaf persons for communication with someone who knows sign language. So far video telephones that were intended to send pictures accompanying voice conversations have been useless for sign language. A whole sequence of signs would be blurred into a single picture because the phones were not designed for real-time updates.

The AT&T Videophone 2500 sends and receives 10 color pictures per second over a standard telephone line. This new telephone incorporates video compression and a 19,200 bps modem into a package that looks much like an ordinary office telephone, except for the tiny built-in color display. Beginning in mid-1992, they will sell for $1500 each and rent for under $30 a day. Ten pictures per second is fast enough to transmit sign language, but only given a large enough picture, sufficient image quality and resolution, and a compression scheme that can handle the fast movements characteristic of sign language (especially finger spelling). Testing with sign language is the only sure way to show whether these phones are suitable for sign language transmission.

"The American video telephone (Picturephone) and the British version (Viewphone) both transmit a picture of the sender to the reader by means of a television raster scan. Unfortunately, Picturephone and Viewphone require a communication bandwidth of [1 MHz, which is 200-300 times the bandwidth available from standard phone lines]. Their enormous bandwidth appetite not only makes them unsuitable for existing telephone transmission and switching facilities, but it makes the development of video telephone facilities economically unattractive."[9] Current research seeks to utilize advanced technology in video compression to develop products which could use existing telephone channels to communicate ASL and finger spelling for persons with hearing impairments.

Progress in image compression has accelerated in the late 1980's and early 1990's. However, as of 1992, it takes many seconds to transmit a clear detailed color picture, with accurate shading and textures, over a standard phone line. Some pictures compress better than others, but using a standard phone line to transmit full-motion color video, at broadcast television quality, is presently beyond the state of the art. It is difficult to predict whether video compression will clear that hurdle by the time phone lines with enough bandwidth for video become cost-effective for individual use, but neither is likely in the next 3 to 5 years. What will happen in 10 years is harder to predict, but for many years, compromises must be made based on technological feasibility. So far, the bandwidth of a standard phone line is too restrictive to transmit such high-quality video.

Transmitting sign language, however, does not require anywhere near that video quality. The human mind can compensate for considerable loss in image fidelity. That compensation may require extra concentration when reading sign language, but many people with hearing impairments would prefer signing over a phone line to typing, especially since the native language of many people who were born deaf is ASL. To them, English is a second language, and they are often more familiar with ASL than English. Extra effort to read sign over a phone line may be preferable to typing on a TDD because signing can be faster and more expressive than typing.

According to Tartter and Knowlton, signers need to see the area in front of the signer's body from the top of the head to waist level, and within a few inches of each side. Signs produced within a 6-inch radius of the chin are most precisely articulated. Battison estimated there are 45 distinct hand shapes, 10 hand orientations, 10 movements, and 25 locations, used in sign language. It could be that facial expressions transmit critical information and head movements may be critical for differentiating affirmatives, negatives and interrogatives.[10]

In the early- to mid-1980's, research on image data compression for sending sign language over phone lines was severely constrained by limited computing power.

Tartter and Knowlton, guided by Sperling, conducted experiments at Bell Labs on communication between two deaf signers. These experiments used 13 light spots on each hand and one spot was placed on the nose. The spots were produced by adjusting the room lighting and using reflective buttons taped to gloves (plus a button on the nose). "The coordinates of the 27 buttons were transmitted with a precision of approximately 1 percent along each axis, at 15 frames per second. Communication was possible, with a data rate of 4800 bits per second," well within the capability of off-the-shelf modems using standard phone lines. However, the system is inconvenient to use and requires training and practice. There are also two serious performance issues: "Finger spelling is difficult, if not impossible," and facial expressions are completely ignored.[11]

Sperling, Pearson, Sosnowski and Hsing followed another approach, transmitting TV-type video, but reducing the resolution of the picture and the number of pictures transmitted per second. Sperling found the minimum bandwidth requirement was 21 kHz (using 30 frames per second, 38 lines per frame, and 50 points per line). That is more than 4 times the bandwidth available from standard phone lines. Pearson found that conversation becomes comfortable at about 100,000 bits per second or more, which is about 7 times the 14,400 bit-per-second limit of modems used on standard phone lines. Conversation becomes impossible below 5,000 bits per second. Sosnowski and Hsing simulated a data rate of 9,600 bits per second, using 8 frames per second. Performance improved with practice, but, based on Pearson's results, that type of system would not permit "comfortable" communication.

It should be noted that no technique for sending sign language over a standard phone line is in general use.

5.3 Commercial Video Teleconferencing

Common carriers, such as AT&T, MCI and Sprint, license video teleconferencing services through other companies, since divestiture does not permit them to deliver the services themselves. These services require the equivalent of many phone lines to transmit the video, however, and that makes them too costly for personal use. For business use, their cost-effectiveness would have to evaluated for individual cases, but they are probably too expensive for day-to-day use in most businesses.

A T1 line has enough bandwidth to transfer 1.544 million bits of information per second. MCI, for example, offers two video teleconferencing options, both of which use video data compression. One uses a full T1 line, the other uses a quarter of a T1 line. The quarter-T1 video teleconferencing service does not update the display frequently enough for sign language using the equipment they provide; the full T1 service is probably fast enough.

However, T1 lines are in demand, both for special services and for carrying telephone conversations. MCI charges about $1850 to $2150 per month for use of a T1 line, plus $9.35 to $5.40 per mile. The base charge increases with increasing mileage; the per-mile charge decreases with increasing mileage. If a per-minute charge is preferred, there is a minimum charge of $500 per month. Some discounts are offered based on usage and time of day, but these charges are in addition to charges for equipment. Equipment can cost $30,000 to $50,000.

These systems are optimized for high image quality, relative to frame rate. A service optimized for much lower image quality at an adequate frame rate would be much more cost-effective for sign language, but only if there is enough demand for it.

5.4 Video Intercoms for Short Distances

When communication is to be over short distances, it may be economical to use simple video equipment and a cable to connect it. At Gallaudet University, for example, a low-cost crib monitoring camera and display are being used as an intercom between two offices. A simple video camera for finding out who is at the front door is another possible source of equipment for this type of innovative application. These approaches can be quite appropriate over short distances and should be publicized, but the cost of video-bandwidth cable becomes cost-prohibitive as distances increase, since it involves a per-foot cost plus an installation cost.

6.0 DEPARTMENT OF EDUCATION'S PRESENT COMMITMENT AND INVESTMENT

The Department of Education's National Institute on Disability and Rehabilitation Research (NIDRR) is currently funding research on transmitting sign language over standard telephone lines and over computer networks. That research is funded as two

Field-Initiated Research Grants at the Department of Computer and Information Science, University of Delaware, Newark.

7.0 ACCESS TO TELECOMMUNICATION SERVICES

Many federal, state, and local laws influence telecommunications for hearing impaired people, just as these laws influence telecommunications for the general population. The most important single law related to telecommunications for hearing impaired people is Public Law 101-336, enacted July 26, 1990. Better known as the Americans with Disabilities Act (ADA), this law has broad implications for all disabled Americans.

Title IV of ADA relates to telecommunications relay services for hearing impaired and speech impaired individuals. It modifies Title II of the Telecommunications Act of 1934 (47 U.S.C. 201 et seq.) by adding Section 225. This section provides that each common carrier providing voice transmission services must also provide telecommunications relay services for hearing-impaired and speech-impaired individuals within three years of enactment of ADA.

Within one year of the enactment of ADA, (i.e., July 26, 1991), the Federal Communications Commission (FCC) must prescribe regulations which:

a) Establish functional requirements and guidelines.
b) Establish minimum standards for service.
c) Require 24 hour per day operation.
d) Require users to pay no more than equivalent voice services.
e) Prohibit refusing calls or limiting length of calls.
f) Prohibit disclosure of relay call content or keeping records of content.
g) Prohibit operators from altering a relayed conversation.

The FCC must ensure that the regulations encourage the use of existing technology and do not discourage or impair the development of improved technology.

The national relay service will probably involve several hundred million calls a year and will be expensive. Any development which shaves a few seconds off an operator's time on a call will mean significant long term monetary savings. This puts tremendous pressure on the telephone industry to develop an efficient technologically advanced service. The cost of sign language transmission equipment for use on standard phone lines is presently high; but relay time may be reduced by offering sign language transmission as an optional substitute for using TDDs. This would require the addition of a sign language interpreting relay service when only one of the two parties is using sign language. No interpreter would be required if both parties use sign language. Common carriers may find it cost-effective to offer reduced rates for lines that have some extra bandwidth for sign language, as a way of saving on the cost of TDD relay services.

It may be advantageous to modify the requirements of divestiture to allow common carriers to provide sign language transmission services if it is not cost-effective for these services to be provided by other companies.

8.0 POTENTIAL ACCESS IMPROVEMENTS WITH ADVANCED VIDEO COMPRESSION TECHNOLOGY

Persons with hearing impairments can potentially benefit from advanced video compression technology because it can make it possible for them to communicate, with friends, relatives, and coworkers in sign language. This is an extremely important advance because hearing impaired persons, especially those who were born deaf, are often accustomed to communicating through sign language. For that reason, the use of English for TDD communication is often difficult and uncomfortable.

Relay services for the deaf would also greatly enhance communication between the hearing and hearing impaired communities if they accommodated the use of sign language in addition to the use of TDDs. Conversation would be potentially much faster and more natural through the use of sign language.

Video services over phone lines could also make it possible to share the resource of sign language interpreters.

Video services may also benefit persons with hearing impairments who rely on lip-reading in combination with hearing speech. Likewise, cued speech, a technique developed at Gallaudet University for providing visual cues with speech for the hearing impaired listener, would benefit from providing video with speech.

9.0 ADVANCED TECHNOLOGIES FOR SIGN LANGUAGE OVER A STANDARD PHONE LINE

Many new applications in video technology will be digital, as is underscored by the Federal Communications Commission's recent statement that it favors an all-digital high-definition television (HDTV) approach. The technologies of digital video and audio integrate the worlds of broadcasting and communication with the world of computing. Ten years from now, these three industries will not be distinguishable.

9.1 Edge Detection

Computer-generated cartoons, produced by a signal processing technique called edge detection, are currently favored as a potential technique for transmitting signs over a telephone line. Current research in this field is being conducted at the University of Delaware, funded by the Department of Education.

A black-and-white video camera delivers a picture of the signer to a computer roughly 10 times every second. Each picture is processed by enhancing lines that are likely to be edges of objects: shoulders, arms, fingers, and some facial features. These edges are then output in the form of a line drawing, throwing away everything else in the picture. The result is a very compact representation of a much more complex image, and it looks just like a cartoon. Image data compression is much more effective on line drawings (cartoons) than on more detailed images, and cartoons are more intuitive to the viewer than reflective buttons or a low-resolution TV image.

Edge detection does impose certain requirements, though. As it is being done at the University of Delaware, edge detection "will most likely detect a dark tie on a light shirt, or detect the pattern of a dress. For this reason, someone using the sign language telephone must wear a dark, solid, non-reflective top and be in front of a dark, solid, non-reflective background. Doing this eliminates unwanted features while providing high-quality contrasts ... ."[12]

If you recall, the use of reflective buttons was much more restrictive, requiring signers to wear special gloves, a button nose, and custom room lighting. The use of very low-resolution video was also more restrictive than edge detection, requiring extremely contrived room lighting and, in at least one case, custom gloves. All of the techniques restrict the backdrop behind the signer.

Based on the limitations imposed by transmission speed, the University of Delaware researchers determined that the feature-extracted images should use no more than 800 pixels, resulting in a frame size of 128 x 128 to 256 x 256 pixels. "While signs themselves were 90% intelligible or better at frame rates as low as 6 frames per second, the quick and subtle movements of finger spelling were 90% intelligible at frame rates of 10 frames per second or above."

"Any kind of real-time image processing takes a tremendous amount of computation power, far more than current personal computers have." Thus, the University of Delaware is using five transputers for this task, with a sixth reserved for expansion.[12] Transputers allow a personal computer to perform parallel processing, which translates to packing the power of several computers into one case.

"The camera used is a black and white CCD camera typically used for security and surveillance purposes. This type of camera gives better images for feature extraction than standard camcorder video cameras. A shutter speed of 1/100 seconds is used to eliminate blur in the captured images. Using this set-up, [they] have been able to achieve a frame rate of 9-12 frames per second with a 256 x 256-pixel resolution."[12]

The image processing required for edge detection is expensive, but that cost will be brought down by the use of application-specific integrated circuits (ASICS) in the next few years. Edge detection has many other applications such as surveillance and robotics, so it is also of interest to the military, for law enforcement, for industrial applications, and eventually, for consumer applications.

The quality of sign language cartoons may also benefit from the use of anti-aliasing, which is a technique for smoothing the jagged, grainy lines of low-resolution images. Anti-aliasing can only help if the display is essentially better than the image it is displaying, but that may often be the case. A small display may be used to make a low-resolution image look better, but it may be worth considering the option of using a larger display with anti-aliasing, to reduce eye strain.

9.2 Fractal Compression

Edge detection is by no means the only algorithm currently available for image compression, as Section 4 of this report attests. The compression schemes described in that section are effective for low to moderate amounts of compression. In combination, they can produce moderately high levels of compression, as shown in Table 1, but sending sign language over a phone line requires an extremely high level of compression. Sending 256 x 256-pixel images at a rate of 10 frames per second requires compression down to 0.0176 data bits per pixel, assuming an asynchronous modem is used at 14,400 bits per second. (An asynchronous modem has a minimum of 20% overhead in its data transmission. Synchronous modems have less overhead but are seldom used for consumer applications. Either way, the compression requirement is extremely high.)

With the possible exception of edge detection, fractal compression is the most promising video compression technology for extremely high compression ratios, given that phone transmission of sign language has a high tolerance for selectively throwing out video information. Fractal compression is based on generating a mathematical representation of aspects of a picture, based on repeated patterns called fractals that often occur in nature (pine cones, for example, are fractals). Fractal compression can require a great deal of processing power for some applications, but it should be investigated for sending sign language over a phone line because it tends to produce hierarchical representations of images. Only certain image details are required to represent sign language intelligibly, and fractals may be helpful in selecting the right details.

Fractal mathematics make it possible to represent parts of an image as a whole, rather than pixel by pixel. "The process begins by taking a digitized image of the subject," as is done for edge detection. "The image is then broken up into segments using image process techniques that include color separation, edge detection and texture variation analysis. The segments are then checked against a library. The library contains relatively compact sets of numbers called iterated function codes that will reproduce the fractals required to develop the image segment. The library is cataloged so shapes that look similar are located closely together. Additionally, nearby codes correspond to nearby fractals. The structure of the catalog permits automated library searches for fractals, which, when combined will approximate the segment. Once the iterated function codes are found for each segment the original digitized image can be discarded and the codes retained, thereby achieving the compression."[13]

Table 1. Compression Schemes

SCHEME	COMPRESSION	DETAILS
Interpolated Differential Pulse Code (IDPCM) Modulation	4:1 to 16:1	Near Lossless at 8:1
Discrete Cosine Transform (DCT)	4:1 to 16:1	Near Lossless at 8:1
Vector Quantization (VQ)	4:1 to 16:1	Near Lossless at 8:1
Joint Photographic Expert Group (JPEG)	20:1 to 30:1	Almost Lossless at 20:1 Poor Quality at 100:1
UVC Corp Proprietary	25:1 to 35:1	30 Frames/Sec
CCITT H.261 (p*64)	50:1	Interframe VHS Quality at 50:1 Lower Quality at 100:1
Wavelet Compression	50:1	50:1 Typical 100:1 Distorted
IDPCM & VQ	72:1	Image Quality Not Available
DCT Interframe	50:1	Good Quality
Prediction/DCT/Motion Compensation	133:1	Lab Model Demonstration at 100:1 Equivalent to MPEG
Moving Pictures Expert Group (MPEG)	100:1	Interframe Predicted VHS Quality at 100:1 Standard Not Final Yet
Fractal Compression Algorithms	40:1 to 160:1	Intraframe 640x400x24 Bit Image Processing Intensive

Two demonstrations of gray-scale fractal compression, in the form of videos from Iterated Systems in Norcross, Georgia, indicate that the technology can perform 40:1 compression of relatively general images. Some of the claims of much higher compression rates that appear in various articles are based on assumptions that are not relevant to this type of application, but fractal compression should at least be tested for this application; it is difficult to predict how well it will work without trying it.

9.3 Sending Sign Language Through Computer Networks

Computer networks are typically designed to carry much more bandwidth than standard telephone lines. Often, computer networks go through the telephone system, but sizeable networks typically use specially installed digital lines, which would be capable of carrying many telephone conversations at once. Local area networks (LANs) also use cables that provide far more bandwidth than a standard telephone line can carry. That extra bandwidth makes it possible for many computer networks to carry sign-language conversations with far less image data compression than would be required over phone lines. In some cases, the frame rate and resolution can be reduced without the need for any other form of image data compression.

Low- to moderate-speed local area networks (LANs), such as Ethernet, reach speeds up to 10-20 million bps. These speeds are even available over wireless (radio or infrared) networks, though wireless networks are generally more expensive than networks that use cables. At the low-speed end, networks designed for only a few users may use less expensive wire such as telephone wire, limiting them to much lower data rates. At the high-speed end of the scale, optical fiber computer networks have transfer rates of 50-150 million bps, but they tend to be used to connect smaller networks rather than individual users. Much higher data rates are available for special applications, but costs limit their size and popularity. Table 2 shows the data rates available with some popular networking schemes.

Table 2. Some Network Protocols and Their Bit-Rate Regimes

SERVICE	BIT-RATE REGIME
Conventional telephone	0.3-56 kb/s
Fundamental bandwidth unit of telephone company (DS-O)	56 kb/s
Integrated-services digital network (ISDN)	64-144 kb/s
Personal computer local area network	30 kb/s
T-1 (multiple of DS-O)	1.5 Mb/s
Ethernet (packet-based local area network)	10 Mb/s
T-3 (multiple of DS-O)	45 Mb/s
Fiber optic ring-based network	100-200 Mb/s

Source: Jurgen, Ronald K. "Digital Video: Putting the Standards to Work," *IEEE Spectrum*, March 1992, pp. 28-30.

The University of Delaware is also involved with researching sign language access through computer networks. They describe several ways in which computer networks can be used to communicate in sign language, analogous to the ways typed (or more recently, voice) messages can be sent over computer networks.

"Talk" utilities allow two network users to type to each other over the network, which is clearly analogous to using TDDs. Typically, half the user's screen will be reserved for what he types and the other half for what the other user types. A similar approach could be used for sign language, but both users would only need to see what the other person is signing. There would be no need for a split screen.

Electronic mail (E-mail) is something like telephone voice mail, but messages are usually typed and sent over the computer network when resources become available. Messages may sometimes be sent right away, but often they do not arrive for several minutes. On a large or very busy network, messages may be held for hours or even days, then sent during a lull in network traffic. A video version of E-mail would also be possible, and it could be useful to anyone, not just the hearing impaired population.

Both of these services would consume network resources because computer networks can only send so much information at one time. The effect of this would be slower network operation, varying with time. Of course, any use of a network slows it down, so the real issue is whether sign language transmission stresses the network more than routine use. The answer depends on factors such as the kind of information being transferred during routine network use, the number of users, the capacity of the network, and the difference between peak loading and typical network load.

E-mail has one distinct advantage over "talk" utilities from this standpoint. E-mail is normally sent more slowly with increasing message size and network loading, but that minimizes its effect on the network. Network loading due to a video "talk" utility could not be spread out over time readily, although there may be ways to conserve network bandwidth when signers pause.

One of the most significant applications of these techniques would be to allow one or more sign language interpreters to be shared by everyone on the network. The only special requirements would be that sign language users would need a special camera, which could be off-the-shelf, at their work stations, and the network would also need special software to handle the sign language features. As is often the case, other uses for sending video over a computer network are starting to emerge, since some ideas are much easier to convey graphically than in words, but they are probably not urgent enough to have a significant effect, in and of themselves, for several years.

Since computer networks often have less severe bandwidth requirements than standard phone lines, their use for sign language is sometimes possible with off-the-shelf equipment and special software. Sometimes, this would make sign language over a computer network much less expensive than sign language over a telephone line.[5] However, the effort involved to set up such systems would probably be greatly reduced by efforts to coordinate the effort between organizations. The alternative would probably be much duplication of effort, and the resulting expense might discourage organizations from implementing sign language on their networks.

"The advantage of using the [computer network] over the sign language telephone is that you can see a real person rather than a line-drawn representation of the person. Certain nuances of signing can be more readily understood by seeing a real person versus a line drawing. Also, such restrictions as having to wear a dark solid top are eliminated by using the [computer network approach]."[12] Of course the computer network approach also requires a computer network that has at least some extra bandwidth. Not every office has a computer network, and an overloaded network may be inadequate to support the addition of video for sign language.

9.4 Advanced Technologies for Commercial Video Teleconferencing

Existing technologies for commercial video teleconferencing are expensive when applied to sign language transmission, but this is largely due to their being optimized for high picture quality rather than high frame rate. This is not so much a technical issue as a cost/demand issue, since developing and fielding a low-image-quality high-frame-rate video teleconferencing system has evidently not been a high priority of companies that offer video teleconferencing services.

Eventually the Broadband Integrated Services Digital Network (B-ISDN) will probably cover all homes in the U.S., providing video bandwidth over ISDN lines at a cost that will make their use popular in homes. That would probably solve, or lay the groundwork for solving, most of the telephone access problems currently faced by the hearing impaired. However, universal availability of B-ISDN requires replacing the existing telephone network with wider-bandwidth transmission lines, such as fiber optics. Completing that upgrade in a few years would be cost-prohibitive.

Narrowband ISDN (N-ISDN), providing capability for 64,000-144,000 bps is already available in some areas and some buildings, however. Increasing network bandwidth should cause wideband channels to slowly become more affordable.

One way to achieve higher data rates over the telephone lines is to use multiple telephone lines. This technique is known as line multiplexing. Line multiplexing is not practical for calls between individuals' homes because most individuals do not have more than one telephone line. However, the technique may be practical for communication among business, Government, and educational institutions. These institutions commonly use public switched networks (PBXs). If a building has wiring, such as computer network, that is capable of handling about 56,000 bps, sign language calls to and from that building could be routed through four modems, all operating at 14,400 bps at a central location in the building. The combined transfer rate of over 56,000 bps would be sufficient to carry sign language with only minimal video data compression. Various tradeoffs are also possible. For example, three 19,200 bps modems could do the work of four 14,400 bps modems, but the slower modems would cost less and be more common. Similarly, reducing the number of modems and/or phone lines would reduce the resulting image quality and/or require more expensive video data compression techniques. Line multiplexing is based on using the same telephone lines for both sign language and for voice conversations, depending on which is needed at the time. Organizational size plays an important part in determining how cost-effective line multiplexing can be for a particular institution, but line multiplexing should be explored as a way to use multiple phone lines for sign language communication.

10.0 COST CONSIDERATIONS OF ADVANCED TECHNOLOGY

The cost of sending sign language over any transmission medium depends on four factors: the cost of the image capture and display equipment, the cost of the computer equipment used to process the video information at either end, the cost of the modulation and demodulation equipment at either end, and the cost of using the transmission medium.

Disregarding schemes that require contrived lighting conditions and special gloves, the dominant cost for sending sign language over standard phone lines is presently that of the computing equipment at the transmitting end, although two-way transmission would require that equipment to be duplicated at both ends. The reason for such high computing costs (many thousands of dollars) is the mathematical complexity of achieving extremely high levels of image data compression through image processing. The computing costs will moderate as application-specific integrated circuits (ASICS) are developed for image processing applications. The cost of the image capture and display equipment can probably be quite reasonable (a few hundred dollars), and the computing equipment required at the receiving end would probably be part of the computing equipment for transmitting at that end. Very high speed modulation and demodulation equipment (modems) will probably be around $200 by the time such systems could be fielded, and the cost of using standard phone lines is well-known.

Commercial video teleconferencing costs depend on all of the system cost factors mentioned plus how well the system is optimized for the task at hand. The cost of lines with wider bandwidth is sure to drop as more telephone network bandwidth is demanded and installed, but the degree of optimization depends on whether effort is concentrated on low-video-quality moderate-frame-rate video teleconferencing. That is very difficult to predict.

Line multiplexing permits a compromise between line costs and transmission equipment and image data compression costs.

Low-cost video intercoms are already cost-effective for short distances, and will become even more attractive as image capture and display technologies improve. Their applicability to distances greater than a room or two depends on the cost of cables and their installation costs. Installation costs may decrease in the next few years as special wiring in homes, and especially offices, becomes more extensive. However, no dramatic changes should be expected. Systems now suitable for short distance are likely to remain so for the near future.

Computer networks have the potential to be a boon for sign language transmission in the near future because most of the cost of incorporating sign language into a moderate-sized network is in the cost of installing the network itself. As computer networks become more and more common, it could become increasingly attractive to include sign language capability in them, given some guidance as to how to go about it.

11.O COST BENEFITS TO PERSONS WITH SENSORY IMPAIRMENTS OF EARLY INCLUSION OF SIGN LANGUAGE CAPABILITY IN TELECOMMUNICATIONS

Consumers with hearing impairments will benefit from early efforts at inclusion of sign language capability through early consideration in important decisions.

As special integrated circuits for implementing edge detection and fractal compression are developed, it is important to be prepared to test them for transmitting sign language over phone lines. Early intervention in the development process could ensure that no design decisions are made that make their application to sign language transmission more difficult or expensive.

Likewise, video telephones should be tested for applicability to sign language transmission as they appear on the market. Technologies intended for use by the general public are often more cost effective than limited market products designed for a special population, but they are not always as effective from a human factors standpoint.

Development of sign language transmission features for computer networks will primarily affect the cost to businesses of installing sign language transmission features on their networks. Persons with sensory impairments will benefit by greater access through that installation.

Development of special equipment for video teleconferencing suitable to sign language transmission will initially only affect businesses, due to the cost. Again, persons with hearing impairments will benefit by improved access.

Early consideration of how to develop sign language relay services would improve the eventual quality and cost of those services, as advance planning always does.

Eventually major telecommunications upgrades, such as the adoption of N-ISDN and B-ISDN, will occur. Early consideration of the requirements imposed on ISDN by sign language capability could greatly affect the cost of sign language transmission equipment that must be able to connect to that system.

12.0 PRESENT GOVERNMENT INVOLVEMENT

The Department of Education's NIDRR is presently supporting research in edge-detection-based (cartoon-like) sign language transmission over standard phone lines at the University of Delaware, Newark. NIDRR is also funding research at the same location on transmitting sign language over computer networks.

13.0 TECHNOLOGY TIMELINE

Edge detection technology for sign language over phone lines shows promise. Research in that field should continue, although a final affordable product will probably have to wait until much of the signal processing required can be done in special-purpose computer hardware. Products are probably at least two or three years away.

Fractal image data compression techniques should also be investigated for sending sign language over phone lines, as there is a chance they could produce a breakthrough. However, it is unlikely that, even with a breakthrough, any product could come out of that research for at least three years.

Commercial video teleconferencing is available now, but the cost is what will determine its applicability. Again, it will be several years before costs come down enough for that to be much of a factor. Efforts to field a low-image-quality high-frame-rate video teleconferencing system might bring that to as little as two years, but that would require interest on the part of companies that would deliver the service. Changes to regulations that would permit the telephone system common carriers to provide the equipment would not necessarily speed the process, but they might make such systems more likely to emerge.

Video intercoms for use over very short distances are already practical, and this will improve slightly over the next few years.

Computer networks show the most promise for sign language transmission in the next two or three years, because the technology and bandwidth is almost there already. Efforts to speed this process have the potential to be very effective. Line multiplexing is already emerging as a way to make wider bandwidth more cost-effective.

Sign language relay systems and interpreter-sharing systems should emerge as techniques to transmit sign language over long distances become available.

Finally, no major telephone system network upgrades will be completed in the next ten years, although the upgrade to B-ISDN will be in progress. That will result in lower transmission costs over video-bandwidth telephone lines, but the costs will probably ease down over the next five to ten years or more.

14.0 PROPOSED ROADMAP

Commercial video teleconferencing, and the use of video compression for all classes of computers, will develop on their own. However, applications that can tolerate lower image fidelity but need extremely high levels of image data compression may develop more slowly.

Specifically, adaptations to provide sign language access through telecommunication systems will tend to develop slowly, so the Department of Education should selectively fund these types of efforts. The Department of Education should continue to support development of edge detection technology specifically for applications to sign language, and should consider joint funding with other government agencies that are also interested in edge detection.

Likewise, fractal image compression should be investigated, specifically looking for applications for sign language transmission. Fractal image compression development at Iterated Systems began with funding through the Defense Advanced Research Projects Agency (DARPA), and consultation with DARPA about possible applications to sign language may be helpful. The possibility of joint funding of experiments may also be considered, although it is very important that the special application of sign language be emphasized. Fractal compression itself will develop on its own.

Commercial video teleconferencing should not receive funding from the Department of Education, because it is profitable in and of itself. However, the Department of Education should consider working with companies that provide video teleconferencing services to develop a standard sign language terminal for use over phone lines, if those companies show interest in fielding such a product.

The Department of Education should publicize possible ways to transmit sign language through direct connections over short distances, but little or no development should be required.

Sign language relay systems and interpreter-sharing systems are still two or three years away, at least, but it is important for the Department of Education to anticipate their development and be prepared; otherwise, they will take much longer to develop. A study could be very helpful in this regard.

Last, but certainly not least, computer networks could start to play a role in transmission of sign language in a year or two, but only if a sufficient number of options are investigated and publicized. Efforts to that end are likely to result in real progress.

15.0 POTENTIAL PROGRAM SCHEDULE

Figure 4 shows a proposed schedule for sign language telecommunications development.

	92	93	94	95	96	97	98
Edge Detection Research	X	X	X				
Fractal System Research		X	X	X			
Commercial Teleconferencing		X	x	X	X		
Computer Network Transmission				X	X	X	X
Sign Language Relay Systems			X	X	X		
Video Phones		X	X				

Figure 4. Proposed Schedule

BIBLIOGRAPHY

1. R. Gonzales and P. Wintz. *Digital Image Processing*, Addison Wesley, Reading, MA, 1977.

2. J. Ziv and A. Lempel. "A Universal Algorithm for Sequential Data Compression," *IEEE Trans. Info. Theory*, Vol. IT-23, No. 3, May 1977.

3. T. Welch. "A Technique for High Performance Data Compression," *Computer*, June 1984.

4. W.K. Pratt. *Digital Image Processing*, John Wiley & Sons, NY, 1978.

M.L. Rhodes, J.F. Quinn, and J. Silvester. "Locally Optimal Run-length Compression applied to CT Images," *IEEE Trans. Medical Imaging*, MI-4,84-90, 1985.

5. National Imagery Transmission Format (NITF) versions 1.1. Defense Intelligence Agency DIA/DC-5C, Wash. D.C., July 1988.

6. A. Rosenfeld and A. Kak. *Digital Picture Processing*, Academic Press, NY, 1976.

A. K. Jain. "A Fast Karhunen-Loeve Transform for Finite Discrete Images," Proc. of the National Electronics Conference, Chicago, 323-328, Oct. 1974.

A. K. Jain, S. Wang, and Y. Liao. "Fast Karhunen-Loeve Transform Data Compression Studies," National Telecommunications Conf., 1976.

7. S. Alexander and S. Rajala. "Image Compression Results Using the LMS Adaptive Algorithm," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. ASSP-33, No. 3, June 1985.

8. Sperling, George. "Video Transmission of American Sign Language and Finger Spelling: Present and Projected Bandwidth Requirements," *IEEE Transactions on Communications*, Vol. COM-29, 12, Dec 1981.

9. Sperling, George. "Bandwidth Requirements for Video Transmission of American Sign Language and Finger Spelling," *Science*, Vol. 210, Nov 14, 1980, pp. 797-799.

10. Tartter, Vivien C. and Kenneth C. Knowlton. "Perception of Sign Language from an Array of 27 Moving Spots," *Nature*, Vol. 289, No. 5799, Feb. 19, 1981, 676-678.

11. Letellier, Philippe, Morton Nadler and Jean-Francois Abramatic. "The Telesign Project," *Proceedings of the IEEE*, Vol. 73, No. 4, Apr. 1985.

BIBLIOGRAPHY

12. Galuska, Scott, Zoran Ivkovich, John Gray, Tim Gove and Dr. Richard Foulds. "Developing Visual Technologies for Deaf Persons in a Work Environment," *Proceeding of the World Congress on Technology*, Wash., D.C., May 1991.

13. Kandebo, Stanley W. "Fractals Research Furthers Digitized Image Compression," *Aviation Week & Space Technology*, Apr. 25, 1988, 91.

PORTABLE POWER SYSTEMS

MARCH 1992

Prepared by

Daniel E. Hinton, Sr., Principal Investigator
and
Rainer Kohler

SCIENCE APPLICATIONS INTERNATIONAL CORPORATION
3701 N. Fairfax Drive, Suite 1001
Arlington, VA 22203
(703) 351-7755

1.0 SCENARIO

Portable Power Systems

2.0 CATEGORY OF IMPAIRMENTS

Persons with Visual and/or Hearing Impairments

3.0 TARGET AUDIENCE

Consumers with Visual and/or Hearing Impairments. This scenario on portable power systems provides a means to disseminate information to consumers with visual and/or hearing impairments. These considerations relate directly to persons with hearing/vision impairments who would benefit from equipment that depends on small portable battery technology, such as sensory enhancing equipment. This evaluation will help the equipment designer choose a battery that is appropriate for their particular application.

Policy makers, including national representatives, Government department heads, and special interest organizations. Policy makers can use this scenario to better understand the issues related to access for persons with visual and/or hearing impairments. This information will also help policy makers to guide research and development efforts in the most advantageous directions.

Researchers and Developers. The R&D community will benefit from this scenario through a better understanding of needs of persons with visual and/or hearing impairments. This scenario was created to assist researchers and developers in evaluating portable power supply options for their future equipment designs.

Manufacturers. Manufacturers will benefit through a better understanding of the potential market size and the existing need for portable power systems in electronic devices used by those with visual and/or hearing impairments.

4.0 THE TECHNOLOGY

Current research in battery technologies is being directed at two areas. The first is improvement (power density, life, cost) of the basic battery which has been around for 70 years. The second area of development is in designing new, small, higher energy density, and less expensive batteries both rechargeable and non-rechargeable.

5.0 STATEMENT OF THE PROBLEM

Today's electronic equipment is smaller and more complex. Portable power supplies that energize these electronic devices must be able to handle the discharge rates and also be economical. Currently there are a wide variety of battery types on the market, each with its own characteristics. The following lists and explains some of the factors that should be considered when selecting a particular battery.

The terminal voltage requirement, as well as the cut-off voltage, of the battery is important. Many electronic devices operate only within a narrow range of voltages and will fail or act unpredictably below the cut-off voltage. Current drain can be long or short in duration, low or high in demand, or any combination of the two. Using a battery which is not designed for high current drain can quickly discharge the battery and can potentially lead to an explosion or electrolyte leakage. Another important factor is operating schedule. Will the device be used continuously or occasionally, and if not in use, will the system recharge the battery? Device accessibility is important if it is not possible or convenient to charge batteries. Temperature range for operation and storage is another important factor. High temperatures can cause batteries to discharge quickly, leak, or even explode. Low temperatures can degrade battery performance. Size and weight of the battery can make a device too large and heavy to be portable. However the smaller and lighter the battery, the more expensive it will be. Proper battery selection can eliminate problems and reduce long term costs.

Technology Application Description

Battery usage can be divided into six types: miniature; portable equipment; starting, lighting, ignition (SLI); vehicle traction; stationary; and load levelling batteries. This report concentrates on only miniature and portable batteries. Typical uses of these are: watches, calculators, medical devices, and small portable electronic devices (see Figure 1).

Primary batteries. Primary batteries are one time use, non-rechargeable, and have been in use since 1866. They are generally low-priced, easy to produce, and provide good performance. Below we review conventional dry batteries such as zinc-carbon and lithium batteries.

1. Zinc-Based

- Zinc-Carbon

The original 1866 zinc-carbon battery design has remained basically unchanged. The standard zinc-carbon or Leclanche* battery has a zinc anode, manganese dioxide/carbon cathode, and an electrolyte of ammonium chloride and zinc chloride dissolved in water. Due to its popularity, the zinc-carbon battery has become economical and reliable for low to medium current drain applications. Almost any shape and size of battery is available with voltages from 1.5V to 500V and are available from many sources. The open circuit voltage (OCV) for a single cell is 1.55-1.74V. This battery performs the best in intermittent use in non-extreme temperature ranges. Zinc-carbon batteries have the worst high discharge rate characteristics; however, for many applications the discharge rate is adequate.

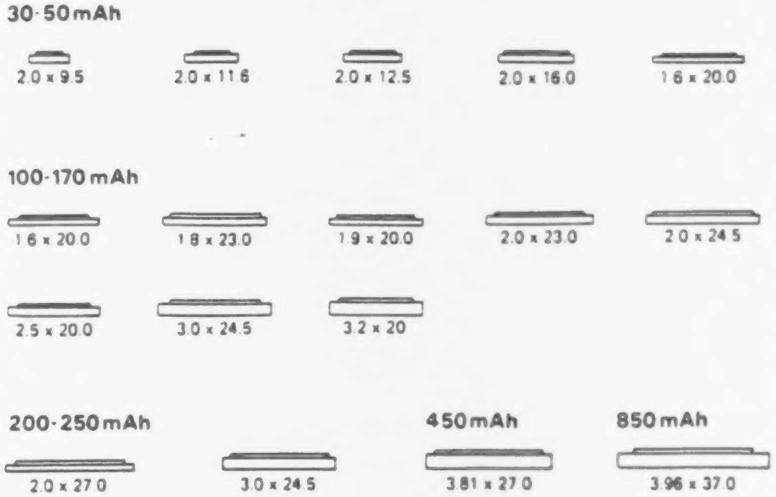

30-50mAh
2.0 x 9.5
2.0 x 11.6
2.0 x 12.5
2.0 x 16.0
1.6 x 20.0
100-170 mAh
1.6 x 20.0
1.8 x 23.0
1.9 x 20.0
2.0 x 23.0
2.0 x 24.5
2.5 x 20.0
3.0 x 24.5
3.2 x 20
200-250mAh
2.0 x 27.0
3.0 x 24.5
450mAh
3.81 x 27.0
850mAh
3.96 x 37.0

Figure 1a. Dimensions (mm) and Capacities of Representative Button Cells

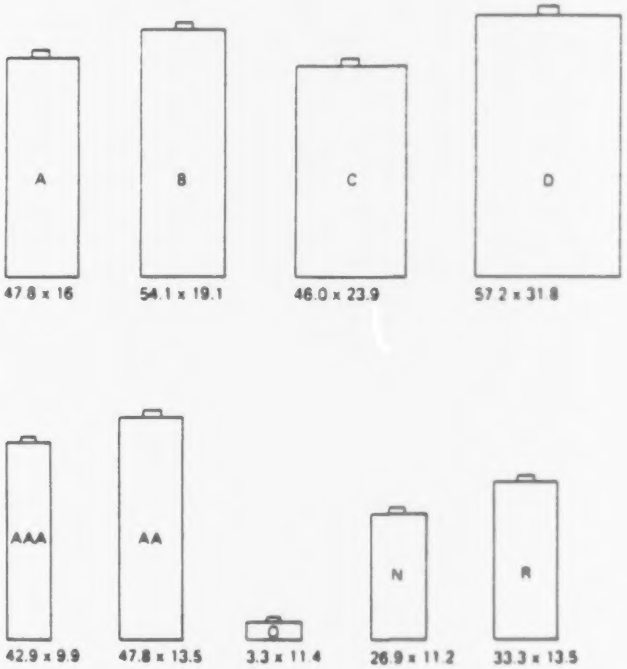

A
47.8 x 16
B
54.1 x 19.1
C
46.0 x 23.9
D
57.2 x 31.8
AAA
42.9 x 9.9
AA
47.8 x 13.5
Q
3.3 x 11.4
N
26.9 x 11.2
R
33.3 x 13.5

Figure 1b. Standard Dimensions (mm) for Cylindrical Primary Cells

BEST COPY AVAILABLE

Leakage of the corrosive electrolyte, which can damage equipment, is possible after heavy discharge. Increased temperatures boost the energy output but lowers the shelf life of the battery. The Leclanche battery quickly deteriorates at temperatures over 52°C (126°F). At temperatures under -18°C (0°F), the zinc-carbon battery becomes close to useless. Storage at 0°C (32°F) helps retain its energy for greater shelf life which is between 2-3 years. The capacity of the zinc-carbon battery depends on the pattern and conditions of discharge more so than other batteries (see Figure 2).

- Zinc-Chloride

The zinc-chloride battery is like the Leclanche battery except it has an electrolyte of only zinc chloride. This electrolyte is very corrosive and requires a better seal to prevent leakage. These batteries are commonly sold under the title of "heavy duty" in retail stores. It is normally constructed with higher quality materials. These two factors increase the cost of each battery, but depending on the application, a zinc-chloride battery may be more economical than the zinc-carbon battery. A single cell has an open-circuit voltage (OCV) of 1.5V. The zinc-chloride electrolyte provides a higher current drain, better low temperature operation, and better performance on continuous drains. The shelf life is also extended to 3+ years. During discharge the battery may dry out from water consumption of the chemical reaction (see Figure 3).

- Alkaline Manganese

Another variant of the zinc-based battery is the alkaline manganese cell. The electrolytes are a combination of potassium hydroxide and potassium zincate. The alkaline battery's constant capacity over a wide range of current drain, wider operating temperature range, lower probability of leakage, single-cell OCV of 1.55V, and shelf life of 4 years make it a good replacement for zinc-carbon and zinc chloride batteries. At high continuous discharge the alkaline battery performs four times better than the two previous batteries; however, for low current drain and intermittent use, the Leclanche battery is more economical to use. The construction of this battery takes on a different physical form and a steel jacket is usually used to contain the highly caustic electrolyte and internal pressure. This results in a higher cost than the zinc-carbon and zinc-chloride battery (see Figure 4 and comparison Figure 5).

- Aluminum/Magnesium Leclanche

Aluminum- and magnesium-based Leclanche batteries replace the zinc anode with either aluminum or magnesium. These batteries have a single-cell OCV of 1.9V. The use of either metal creates two problems: increased corrosion rate and voltage delay. Both metals react with the electrolyte, resulting in anode corrosion and reduced capacity. To increase shelf life, the anode is coated with a film. When

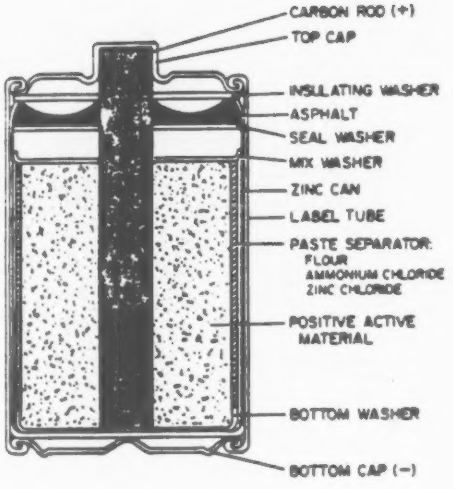
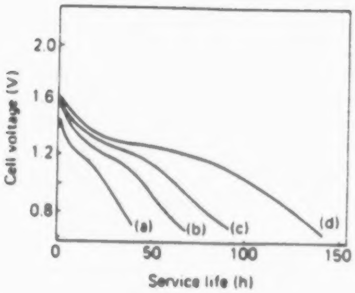

CARBON ROD (+)
TOP CAP
INSULATING WASHER
ASPHALT
SEAL WASHER
MIX WASHER
ZINC CAN
LABEL TUBE
PASTE SEPARATOR:
FLOUR
AMMONIUM CHLORIDE
ZINC CHLORIDE
POSITIVE ACTIVE MATERIAL
BOTTOM WASHER
BOTTOM CAP (−)

Figure 2a. Cut-Away View of Paste-Wall Construction Leclanche Cell



(a) Initial drain: 150mA
(b) Initial drain: 100mA
(c) Initial drain: 75mA
(d) Initial drain: 50mA

Figure 2b. Effect of Discharge Rate on Service Life of D-Size Leclanche Cells Discharged at 2 Hours Per Day

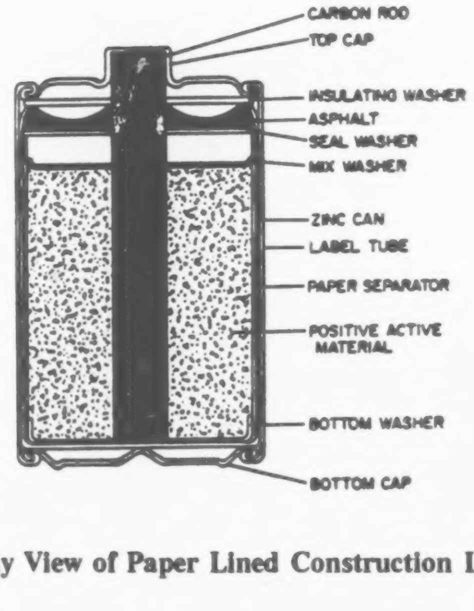

CARBON ROD
TOP CAP
INSULATING WASHER
ASPHALT
SEAL WASHER
MIX WASHER
ZINC CAN
LABEL TUBE
PAPER SEPARATOR
POSITIVE ACTIVE MATERIAL
BOTTOM WASHER
BOTTOM CAP

Figure 2c. Cut-Away View of Paper Lined Construction Leclanche Cell

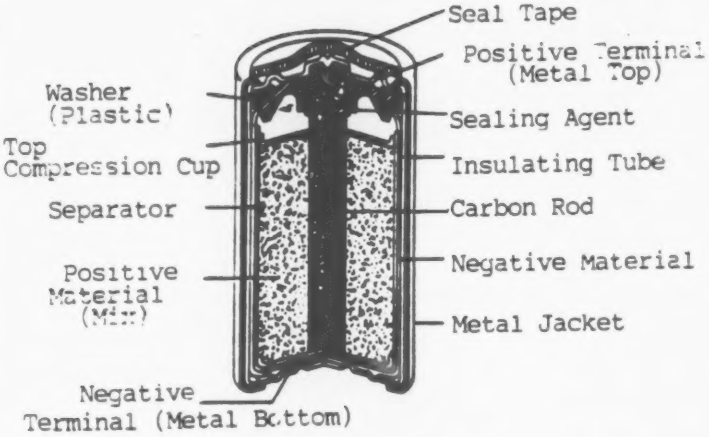

Seal Tape
Positive Terminal (Metal Top)
Washer (Plastic)
Sealing Agent
Top Compression Cup
Insulating Tube
Separator
Carbon Rod
Positive Material (Mix)
Negative Material
Metal Jacket
Negative Terminal (Metal Bottom)

Figure 3a. Cross Section of a Zinc Chloride (Cell)

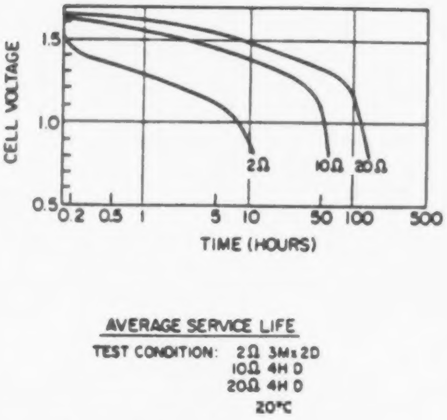


Figure 3b. Typical Discharge Curves of D-Size $ZnCl_2$ Cells Discharged at Various Load and Test Regimes

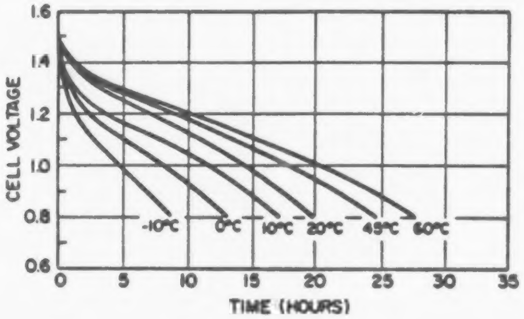


Figure 3c. Typical Voltage Characteristics of a D-Size Zinc Chloride Cell at Various Temperatures, Discharged Continuously on 4 ohm Load

Temperature °C	Normalized Capacity	
	Leclanché	$ZnCl_2$
37.8	1.40	1.15
26.7	1.10	1.05
21.1	1.00	1.00
15.6	0.90	0.95
4.4	0.70	0.85
-6.7	0.45	0.70
-17.8	0.25	0.45

Figure 3d. Effect of Temperature on the Capacity of D-Size Leclanche and $ZnCl_2$ Cells When Discharged Continuously Through 2.25 Ω to a Cut-Off Voltage of 0.9V

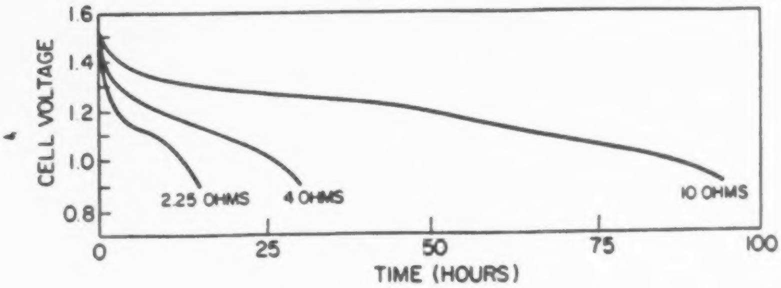


Figure 4a. Typical Discharge Curves for D-Size Alkaline Cell
(Courtesy of Duracell)

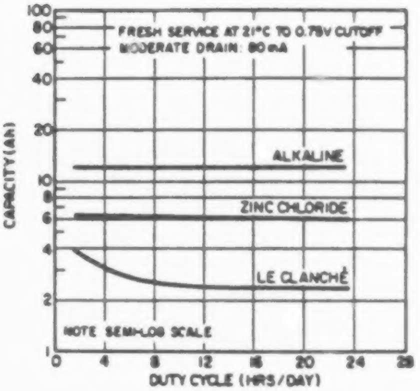


Figure 4b. Variation in Output Capacity as a Function of Duty Cycle
(Courtesy of Eveready Battery Co.)

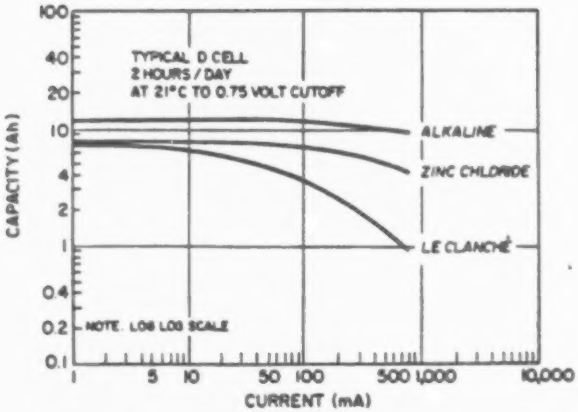


Figure 4c. Variation in Output Capacity

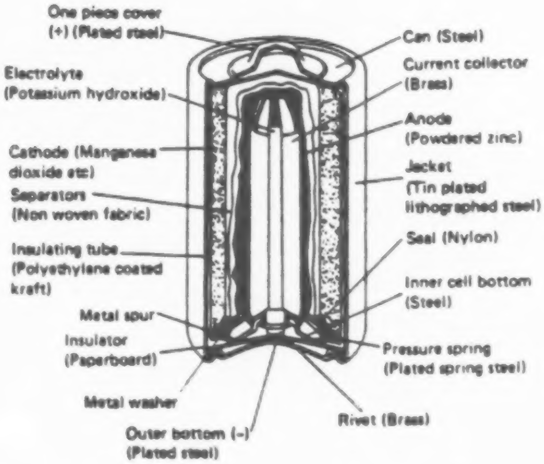
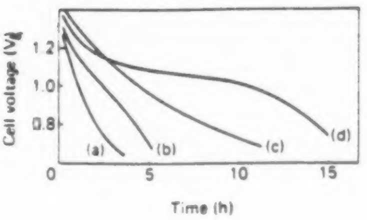

One piece cover (+) (Plated steel)
Can (Steel)
Electrolyte (Potassium hydroxide)
Current collector (Brass)
Anode (Powdered zinc)
Cathode (Manganese dioxide etc)
Jacket (Tin plated lithographed steel)
Separators (Non woven fabric)
Seal (Nylon)
Insulating tube (Polyethylene coated kraft)
Inner cell bottom (Steel)
Metal spur
Insulator (Paperboard)
Pressure spring (Plated spring steel)
Metal washer
Rivet (Brass)
Outer bottom (-) (Plated steel)

Figure 4d. Cross Section of a D-Size Alkaline Cell



(a) Standard Leclanche cell based on natural ore
(b) "High power" Leclanche cell based on electrolytic MnO_2
(c) Zinc chloride cell

(d) Alkaline manganese cell

Figure 5. Comparison of the Performance of $Zn\text{-}MnO_2$ Primary Systems Under 2.25 ohm Continuous Test

discharge takes place, the film breaks down and the battery resumes normal voltage output. This film breakdown time causes a voltage delay to appear at the terminals. Venting is also required to release gas build-up from the dissolved film. Once the film has been destroyed, the cell should be used continuously. Intermittent use requires a film to build up after each discharge on the anode to prevent corrosion. This type of battery has been used mainly for military applications where it is being replaced by lithium based batteries.

- Mercuric Oxide/Silver Oxide

Zinc-mercuric oxide and zinc-silver oxide based batteries are manufactured into miniature or button cells. They contain a zinc anode, an electrolyte solution of potassium hydroxide with zinc oxide, and a mercuric oxide or silver oxide cathode. One of their advantages over Leclanche batteries is their high energy-to-weight ratio, but these batteries do not have the capacity to power high-current applications.

- Zinc-Mercuric Oxide

Zinc-mercuric oxide cells have a very flat voltage discharge curve which is useful in applications that require long steady periods of discharge. Flat low current drain discharge characteristics occur for 97% of operational time. The cell construction provides venting in the case of short circuits or reverse currents, and has an absorbent material to prevent electrolyte leakage. Because of its high energy-to-weight ratio, the zinc-mercuric oxide battery is a third the size of a zinc-carbon battery of equivalent energy capacity and has a 5-year shelf life. At high temperatures the battery performs quite well; however the battery's low temperature characteristics are not very good. Discharge characteristics are relatively independent of load over a wide range of loads with a single cell OCV 1.357V (see Figure 6).

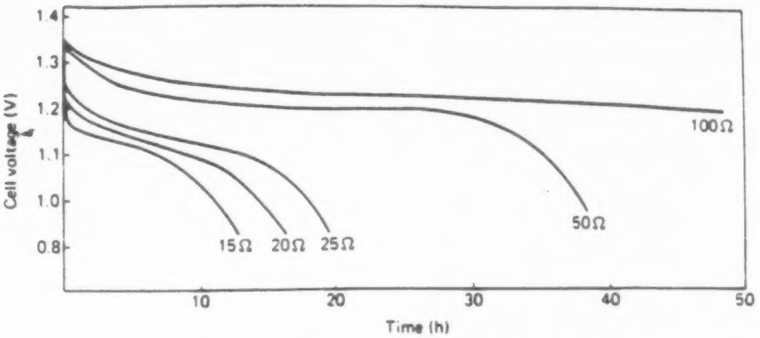


Figure 6a. Discharge Characteristics of 1 Ah Zinc-Mercuric Oxide Button Cell Under Continuous Load at Room Temperature

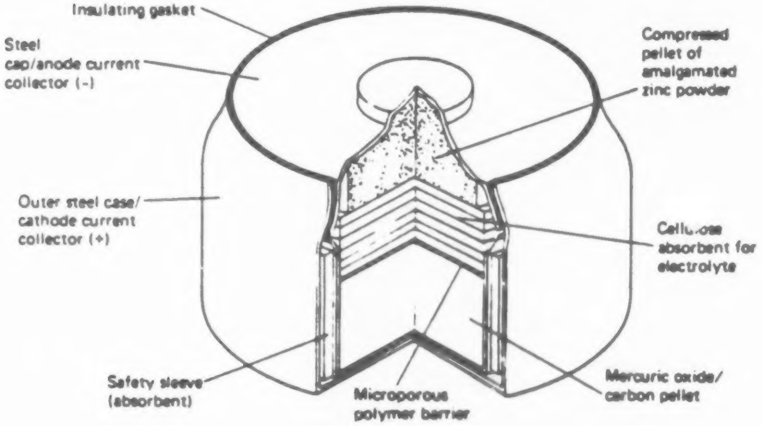

Insulating gasket
Steel cap/anode current collector (-)
Compressed pellet of amalgamated zinc powder
Outer steel case/ cathode current collector (+)
Cellulose absorbent for electrolyte
Safety sleeve (absorbent)
Microporous polymer barrier
Mercuric oxide/ carbon pellet

Figure 6b. Cross-Section of a Typical Zinc-Mercuric Oxide Button Cell

- Zinc-Silver Oxide

Zinc-silver oxide cells have many of the same features as zinc-mercuric oxide. They are half the size of comparable zinc-carbon batteries and have a higher single cell OCV (about 1.6V) than zinc-mercuric oxide batteries. Characteristics at lower temperatures are better then zinc-mercuric oxide. Applications such as LCD displays require high resistance batteries and low current drains, in which case a potassium hydroxide electrolyte would be used. Zinc-silver oxide batteries are more expensive then zinc-mercuric oxide batteries (see Figure 7).

- Zinc-Air

Zinc-air cells have the highest energy density of commercially available batteries because one of their reactants is air. After their first use, however, their shelf life is limited, so they are best suited to frequently used devices, such as hearing aids. Zinc-air cells have a single-cell OCV of about 1.65V, which is similar to most zinc-based batteries, but they have more than twice the energy density of

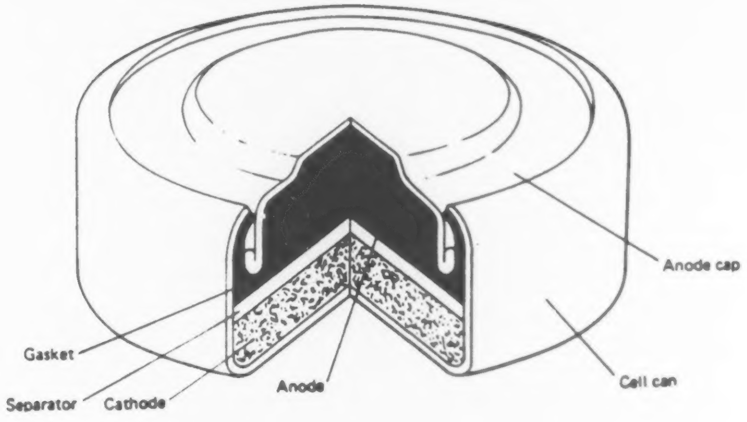


Figure 7a. Cutaway View of a Typical Zinc-Silver Oxide Button Cell (Courtesy of Union Carbide)

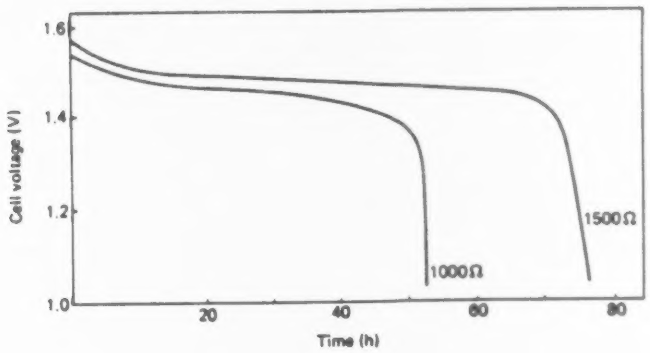


Figure 7b. Discharge Characteristics of 75 mAh Zinc-Silver Oxide Hearing Aid Cell Under Continuous Load at Room Temperature

other commercially available batteries, both by weight and by volume. They can cost twice as much as a mercury battery of the same size, but the difference in battery life can make them more economical. Zinc-air cells have about the same operating temperature limitations as carbon-zinc cells, but their range is shifted up by 5°C (9°F). They are available as button cells and larger units.

2. Lithium-Based

In the last few years, the need for a more powerful and smaller power supply has increased. Conventional dry batteries such as the zinc-carbon have reached their technological peak and mercury/silver based batteries do not meet the required power levels. Research, along with advances in material handling and processing, has produced the lithium anode-based battery. Energy densities of up to three times that of mercury- and alkaline-based batteries, and volumetric energy densities of 50 and 100% higher, have been achieved. Lithium batteries are divided into three groups: solid cathode, soluble cathode, and liquid cathode.

- Solid Cathode

Lithium solid cathodes are divided into four subgroups: polycarbon fluoride, oxosalts, oxides, and sulfides. Each group has its own characteristics as outlined below.

-- Polycarbon-fluorides-cathode batteries typically have a single-cell OCV of 2.8-3.3V. These systems experience voltage delays and have a shelf life of 5+ years. The discharge curve is relatively flat for much of the operating time (see Figure 8).

-- Oxosalts cathode batteries have a single cell OCV of 3.5V. A voltage delay is present at high current drains. A unique characteristic of this system is the two-plateau voltage level. The voltage is constant at one plateau for 75% of the discharge time. At the end of this time the voltage drops to the second plateau where the typical cutoff voltage is 2.5V. Battery replacement is usually done at this time. Some minor swelling occurs during discharge. This system has very good reliability under low continuous drain. Storage effects on battery capacity are less than 1% a year (Figure 9).

-- Oxides cathode batteries have a single-cell OCV of 1.5V or 3.0-3.4V depending on the type of cell. These systems have a wide operating temperature range for moderate to low current drains. At low discharge rates, oxides are as safe as carbon-zinc batteries; however, there is insufficient data at higher operating temperatures. Storage life of 20 years at 21°C (70°F) is typical (see Figure 10).

-- Sulfides cathode batteries have a single-cell OCV of 2.2V. Like the oxosalts, these batteries have a two plateau voltage discharge. The first stage is 2.2V for 80% of its operating time. The second stage is 1.75V, at which time replacement is recommended. Since metal sulfides are good conductors, it is not necessary to use carbon in the cathodes which results in a lighter battery (see Figure 11).

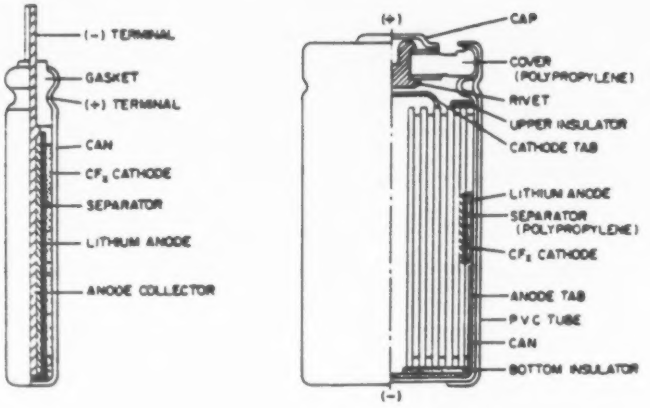

(−) TERMINAL
GASKET
(+) TERMINAL
CAN
CF_x CATHODE
SEPARATOR
LITHIUM ANODE
ANODE COLLECTOR
(+)
CAP
COVER (POLYPROPYLENE)
RIVET
UPPER INSULATOR
CATHODE TAB
LITHIUM ANODE
SEPARATOR (POLYPROPYLENE)
CF_x CATHODE
ANODE TAB
PVC TUBE
CAN
BOTTOM INSULATOR
(−)

Figure 8a. Cross Section of Li/CF Pin and Cylindrical-Type Cells
(Courtesy of Matsushita)

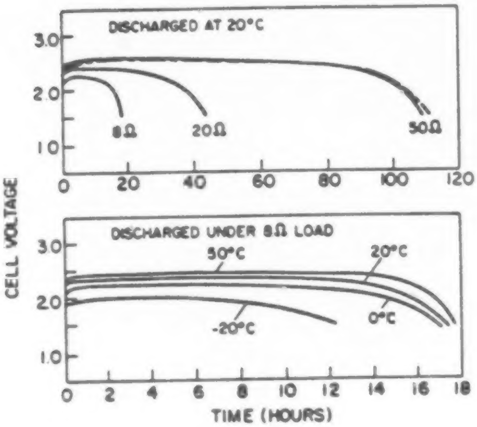


Figure 8b. Effect of Rate and Temperature on Discharge of
Spiral-Wound C-Size Li/CF_x Cell

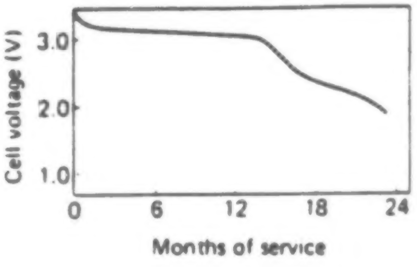


Figure 9. Discharge Characteristics of a Lithium-Silver Chromate Pacemaker Cell of
Nominal Capacity 600 mAh (SAFT Gipelec Li210) Under a Load of 75 kΩ
(Courtesy of SAFT Gipelec)

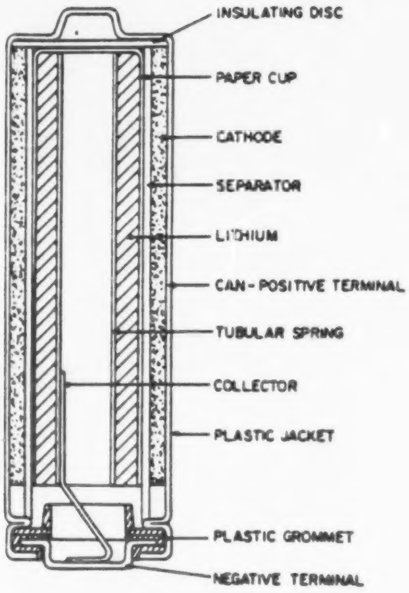

INSULATING DISC
PAPER CUP
CATHODE
SEPARATOR
LITHIUM
CAN - POSITIVE TERMINAL
TUBULAR SPRING
COLLECTOR
PLASTIC JACKET
PLASTIC GROMMET
NEGATIVE TERMINAL

Figure 10a. Construction of AA-Size Li/CuO Bobbin Cell

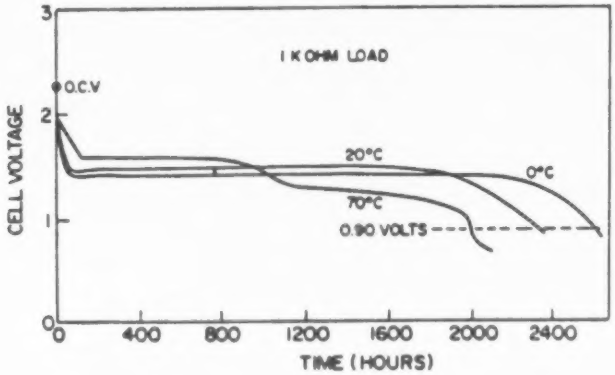


Figure 10b. Effects of Temperature on Continuous Discharge at 1000 ohms for an AA-Size Li/CuO Cell of Bobbin-Type Construction

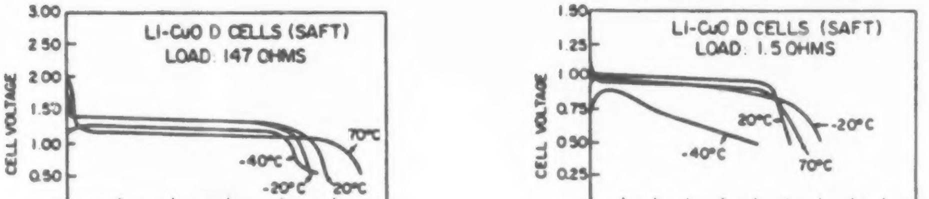


Figure 10c. Effects of Discharge Rate and Temperature on Spiral-Wound D-Size Li/CuO Cells

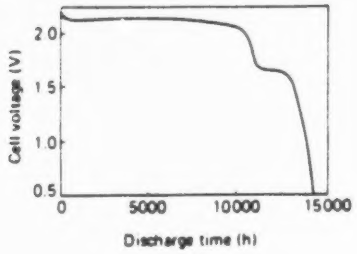


Figure 11. Discharge Curve of a Lithium-Cupric Sulfide Pacemaker Cell at 37°C Under a Load of 12.3kΩ

- Soluble Cathode

Lithium soluble cathode battery systems are all based on sulphur dioxide. Typical single cell OCV is 3.0V with an operating temperature range of -54°C (-65°F) to 70°C (158°F). This battery is good for applications requiring heavy duty power. Low temperature discharge is quite good and pressure within the cell decreases during discharge. At 100°C (212°F), pressure in the cell increases rapidly to a dangerous level. Discharge characteristics are flat with good voltage regulation. A small voltage delay is present. Cells may explode or vent toxic gas in the event of prolonged short circuits or high temperatures. Shelf life is around 5 years (see Figure 12).

- Liquid Cathode

Lithium liquid cathode battery systems have a single-cell OCV of 3.6V. Operating temperatures are -40°C to 75°C, with some voltage delay below -30°C. The battery has an extremely flat discharge rate for 90% of the time and has good high current discharge properties. Special sizes and terminals are manufactured for direct placement on circuit boards. The cells must be hermetically sealed for safety. Explosion may occur at high rates or due to a short circuit or forced discharge. The shelf life of a liquid cathode cell is 6+ years (see Figure 13).

Secondary batteries. Secondary batteries can be recharged repeatedly during their lifetime. They are more expensive than primary batteries and also require a charger.

- Lead-Acid

Lead-acid batteries have porous lead anodes, lead dioxide cathodes, and sulfuric acid electrolytes. Water is formed from the sulfuric acid during discharge and results in a decreased concentration of sulfuric acid. At full charge the concentration is 40% and at full discharge it is 16%. The single- cell OCV is 2.15V

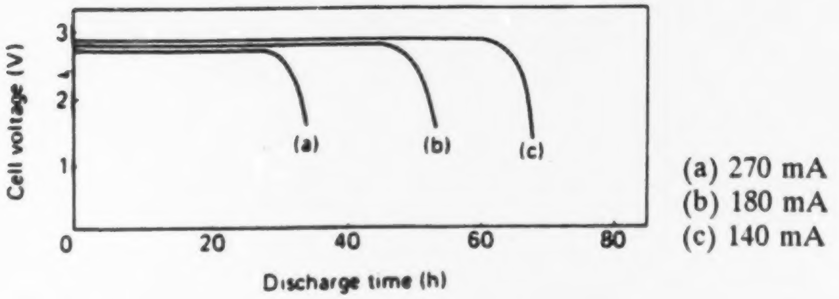


Figure 12a. Discharge Curves for D-Size Lithium-Sulphur Dioxide Cells at Ambient Temperature

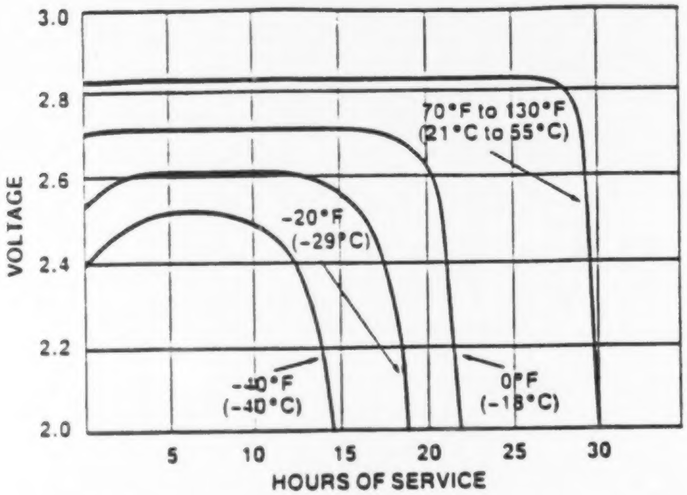


Figure 12b. Temperature Service Effects in Li-SO_2 Cell

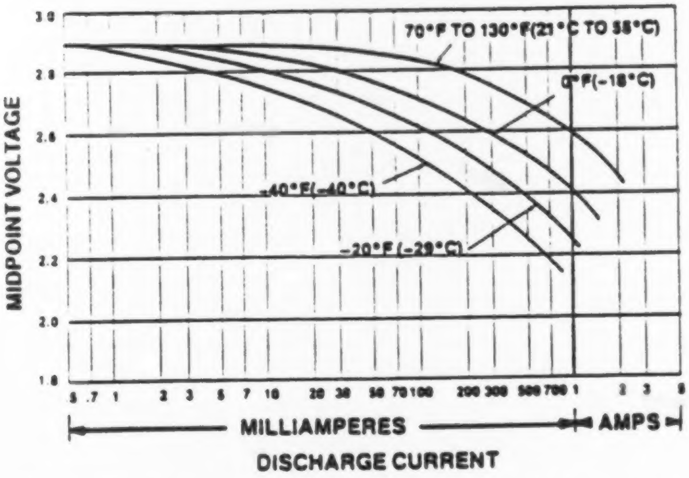


Figure 12c. Temperature Load Effects in Li-SO_2 Cell

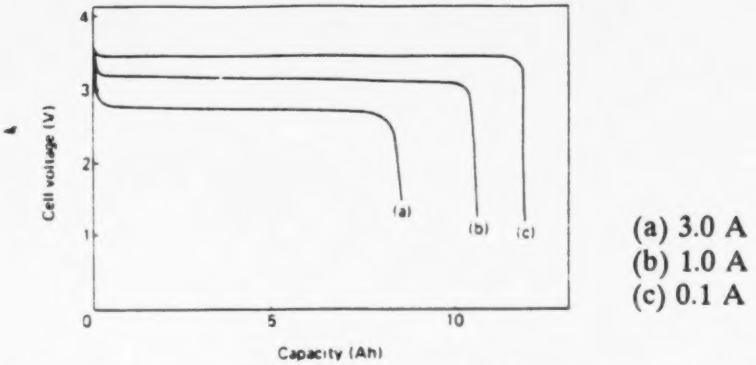


Figure 13a. Discharge Curves of D-Size Lithium-Thionyl Chloride Cells at Ambient Temperature

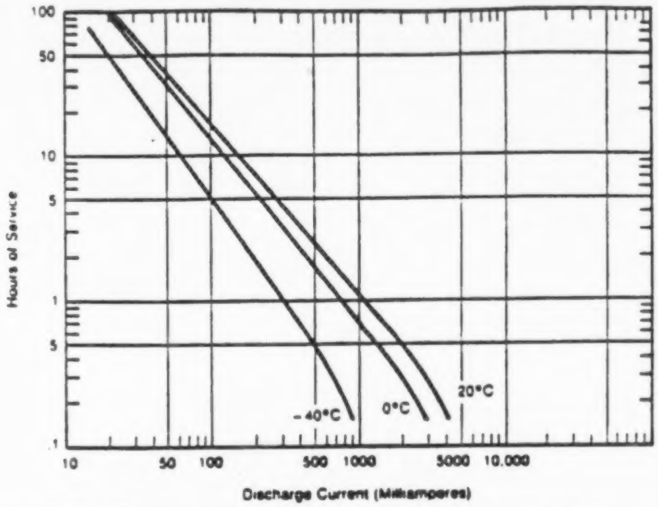


Figure 13b. Temperature Service Effects in Li-SOC1$_2$ Cells

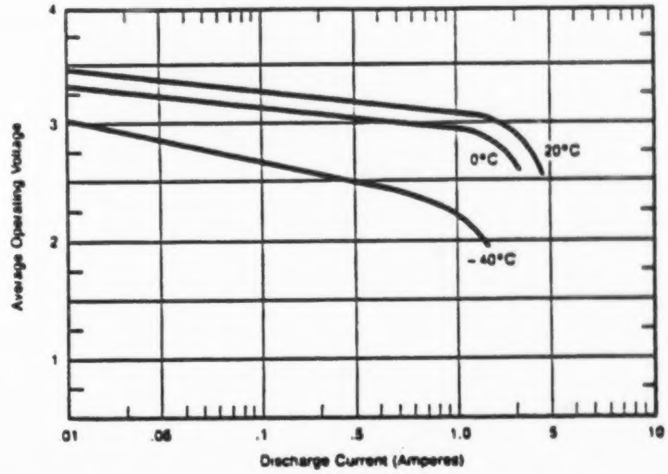


Figure 13c. Temperature Load Effects in Li-SOC1$_2$ Cells

at full charge and 1.98V at full discharge. Insoluble lead sulphate forms on the electrodes during discharge and reduces their capacity, but solutions to this problem have been found. Temperature has an important effect on the battery capacity. Below 0°C (32°F) capacity drops off very quickly. Self-discharge is dependent on temperature, concentration, electrolyte, and most important, on the purity of materials. Battery damage can occur if the battery is left for long periods in a discharged state, operated at high temperature, or at high acid concentration.

Most manufactured lead-acid batteries have safety vents to release internal pressure. However this limits the orientation of the battery. Semi-sealed lead-acid batteries can operate in any position and have been used in many applications, including computers, portable TV's, etc. (see Figure 14).

- Ni-Cad

Nickel-cadmium (Ni-Cad) batteries account for 82% of secondary batteries sold. Because of their popularity, these batteries can be found in many shapes and sizes, and in either sealed or vented construction. System characteristics such as long life, low maintenance, and high reliability make the Ni-Cad an attractive alternative. These batteries use nickel hydroxide for the cathode, cadmium/iron for the anode, and potassium hydroxide for the electrolyte. Sealed cells are designed so that internal pressure is never a problem. By adding additional materials to the electrolyte, a lower and higher temperature operating range can be achieved. Recharge of 80% capacity is typical in most systems. The charge is built up by using a constant current charger. Charging batteries should receive 120-140% of charge to maximize capacity. Excessive overcharging reduces water, increases internal heat, and increases internal pressure which can damage the battery. Discharge characteristics include a flat discharge rate even at high rates. Capacity of the Ni-Cad drops rapidly at temperatures under 0°C (32°F) and above 60°C (140°F). Shelf life is not very good, with losses in capacity of 30-40% in 6 months. The Ni-Cad battery develops a memory after repeated identical low rate discharge. When trying to discharge the battery in a different pattern or past the original level, the amount of available charge is lower, even if the energy is present. Very slow discharge can be used to "erase" the memory and regain full capacity. Long-term storage of Ni-Cad batteries at any charge will not damage the battery (see Figure 15).

- Nickel-Hydride

Nickel-hydride batteries are recent entries into the battery marketplace and offer twice the storage capacity of NiCad batteries at the same weight, without the "memory effect" which limits NiCad capacity. The new batteries, manufactured by Toshiba, Inc., use a nickel-metal hydride formulation based on the electrochemical properties of a hydrogen-absorbing alloy. The process is as follows:

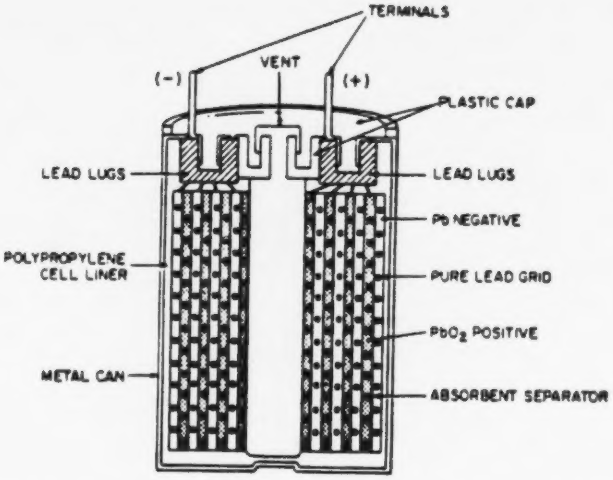

TERMINALS
VENT
(-)
(+)
PLASTIC CAP
LEAD LUGS
LEAD LUGS
Pb NEGATIVE
POLYPROPYLENE CELL LINER
PURE LEAD GRID
PbO₂ POSITIVE
METAL CAN
ABSORBENT SEPARATOR

Figure 14a. Schematic of Sealed Cylindrical Lead Acid Cells

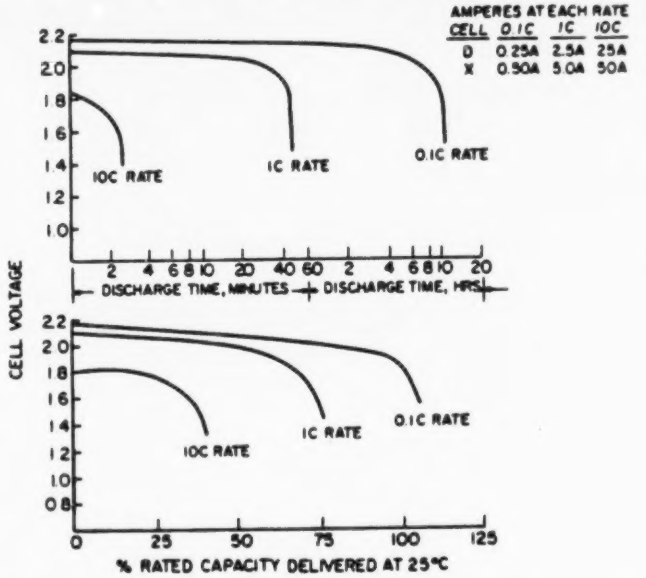


Figure 14b. Typical Sealed Cell Discharge Voltage Time Curves at 25°C

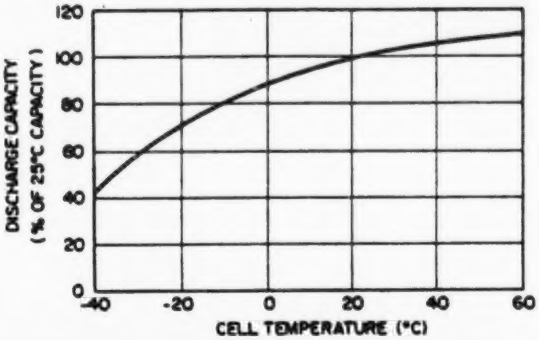


Figure 14c. Typical Discharge Capacity as a Function of Cell Temperature

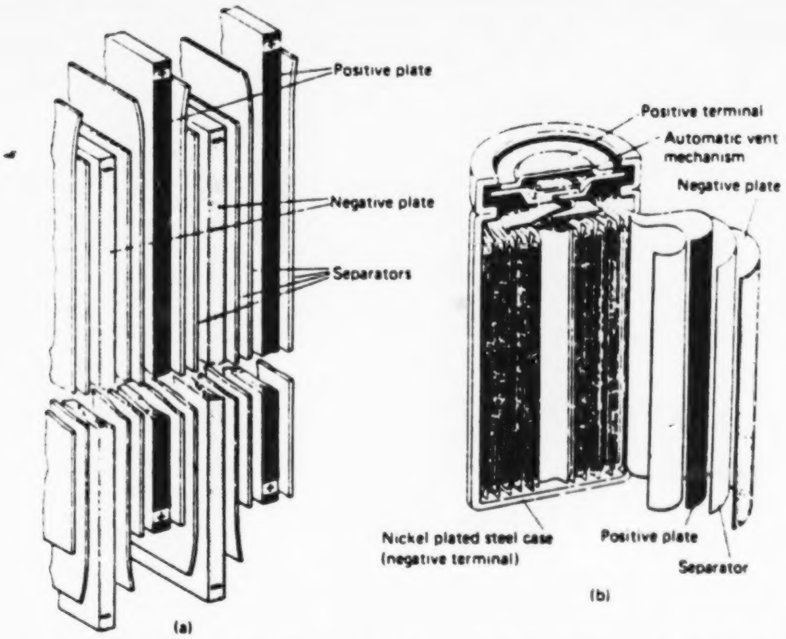

Positive plate
Negative plate
Separators
(a)
Positive terminal
Automatic vent mechanism
Negative plate
Nickel plated steel case (negative terminal)
Positive plate
Separator
(b)

Figure 15. Schematic Cross-Section of Spiral Wound Cylindrical Sealed Nickel-Cadmium

The hydrogen-absorbing alloy becomes metal hydride during the charging of the battery by absorbing relatively large volumes of hydrogen. The reaction is reversed in the discharge cycle. The operating voltage of this battery is 1.2 volts and is compatible with NiCads in its application.

Like NiCad batteries, nickel hydride batteries are completely enclosed with positive and negative hydride batteries are completely enclosed with positive and negative electrodes sealed with a liquid electrolyte. The capacity of the positive electrode determines the useful capacity of the battery. Unlike NiCad, the charge capacity of the negative pole, the hydrogen-absorbing alloy, is far higher than that of the positive pole material which means a much larger quantity of positive pole material can be contained within the structure allowing a significant increase in the volume of the positive electrode and therefore a significant increase in charge capacity.

Nickel-hydride batteries cannot deliver as great a current as NiCad batteries can, but they have 50 percent longer life than NiCad batteries of comparable size and weight.

- Zinc-Silver

Zinc-silver oxide batteries, implemented with a different design than primary zinc-silver batteries, can be repeatedly recharged. These cells are expensive and have poor recharge cycle life. Discharge performance falls sharply--below 10°C (50°F). At higher temperatures the battery can sustain high discharge rates and has an energy density of five times that of a Ni-Cad battery. Like the primary battery version, the zinc-silver oxide battery has two voltage plateaus: 1.7V and 1.5V. At high discharge rates the first plateau disappears and continued high rates increases internal battery temperature. Recharging over 2.0V can cause physical damage to the internal materials. This type of battery can be found mainly in military and aerospace applications (see Figure 16).

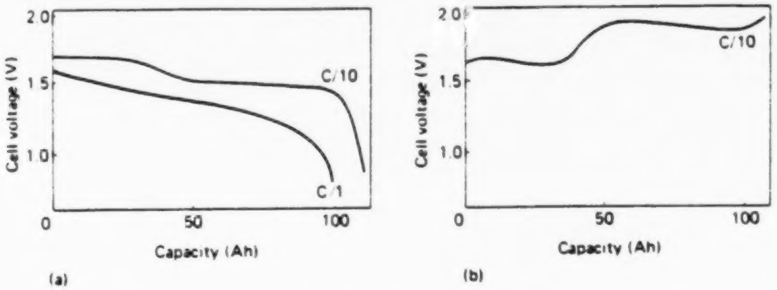


Figure 16. Discharge (a) and Charge (b) Characteristics of a Typical 100 Ah Zinc-Silver Oxide Cell

- Conductive Polymer

Conductive Polymer Batteries are recently developed batteries based on electrically conductive polymers. Electrically conductive polymers are opening new markets for plastics normally considered insulating materials. For battery applications, polymers are generally used as the cathode, replacing heavy metal electrodes and as a solid electrolyte replacing volatile liquid electrolyte solutions.

Bridgestone Seiko, Japan manufactures a three volt rechargeable button-cell battery using a conductive polymer, called polyaniline, as: the cathode. This battery contains a lithium - aluminum - alloy anode and a solution of $LiBF_4$ for the electrolyte. This battery operates as follows: as voltage is applied during the charging cycle, polyaniline is oxidized, removing electrons, leaving holes (positive charges). $BF4^-$ anions from the electrolyte enter (dope) the polymer making it conductive. The electrons removed from polyaniline travel through the external circuit, causing Li+ ions from the electrolyte to deposit as lithium metal at the anode.

The reverse discharge reaction is spontaneous. Electrons from the lithium anode travel through the external circuit and are harvested outside the battery, as lithium dissolves back into solution at the anode. Completing the circuit, electrons enter the polyaniline, reducing (dedoping) it to the insulating form as $BF4^-$ anion is ejected.

The benefits of the conductive polymer cathode are lighter weight and high energy density (40 W-Hr/Kg, comparable to NiCd batteries). The battery is characterized by a long cycle life however, has a short shelf life.

A new ionically conductive polymer electrolyte battery, called ultracell, offers high safety, ease of fabrication and design flexibility, with an energy density 5 to 6 times higher than NiCd batteries. Ultracell is manufactured by a Danish-American joint venture in San Jose, CA and is a rechargeable lithium battery.

Lithium is a desirable electrode material because of its light weight and high electropositivity, which translate to high energy density. Lithium electrodes, however, have a tendency to grow fine filaments into liquid electrolytes which can short circuit the battery. The heat generated can cause explosions or venting of electrolyte solvent, creating a fire hazard. Solid conductive polymer electrolytes present a physical barrier to filament growth and eliminate the use of potentially hazardous liquid electrolytes altogether.

The Ultracell fabrication system is also revolutionary. Rather than "building" the battery, as shown in Figure 17, it is made by laying down thin films of electrode/electrolyte/electrode/current collector, resulting in continuous manufacture of 3-volt cells 100 to 150 m thick. The battery can be folded, rolled, or cut and stacked to meet design requirements.

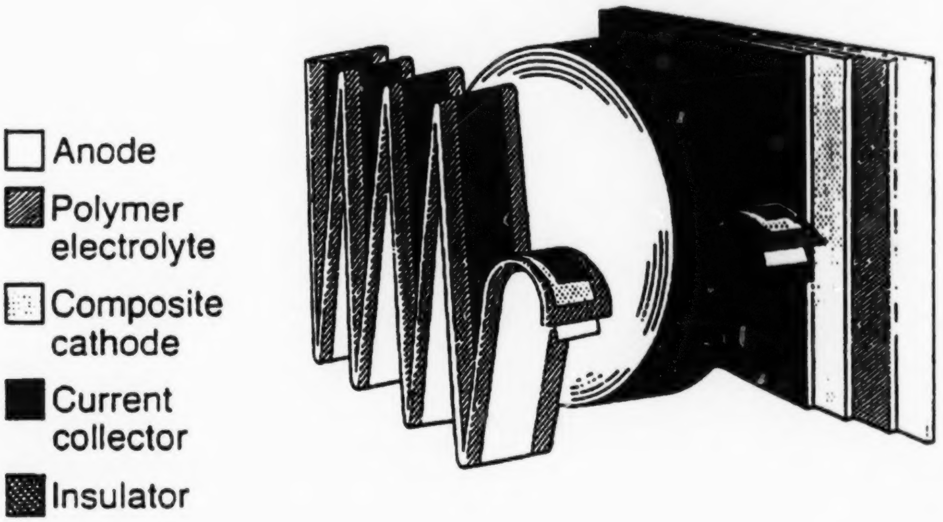


Figure 17. Ultracell Experimental Cell Configuration

Because the ionically conductive polymer electrolyte is not as conductive as liquid electrolytes, these batteries do not discharge rapidly, and are not suggested for generating high surges or pulses of power. Their forte is steady output, long life, and high reliability for cellular phones, portable computers, timepieces, calculators, etc.

Overview. Table 1 summarizes the options currently available for portable power, listed in order of theoretical capacity. It is based on the table on pages 32-33 of *IEEE Spectrum*, March 1988.

6.0 DEPARTMENT OF EDUCATION'S PRESENT COMMITMENT AND INVESTMENT

The Department of Education has not been involved in funding battery supply and development. No involvement is expected in the future except as part of development efforts to meet form, fit and function of devices for persons with visual and hearing impairments.

7.0 ACCESS TO BATTERY TECHNOLOGY

No special barriers to portable power technology access are foreseen other than cost and standardization, which affect all consumers to some extent. Batteries of at least flashlight size have significant differences between the few commonly used sizes and shapes; however, the standard sizes and ratings of smaller batteries, used for hearing aids, watches, etc., fill catalogs. To some extent, that is because smaller batteries require more compromises than larger ones, but equipment manufacturers also want to sell their own batteries: High-performance batteries often require frequent replacement, and that can be profitable. If equipment manufacturers could reach a compromise and produce fewer small standard batteries, it would encourage competition and lower prices, so all consumers would come out ahead, especially, people with impairments that make them dependent on battery-powered technologies.

8.0 POTENTIAL ACCESS IMPROVEMENTS WITH BATTERY TECHNOLOGY

Battery technology should be accessible to persons with both hearing and vision impairments. Care should be taken in the design of batteries to allow easy installation and replacement. The Department of Education should pay particular attention to types and physical accessibility of the batteries and battery compartments in all designs built under their auspices. Specifically, it should be easy for a visually impaired person to identify the proper batteries to use in an assistive device, locate and open the battery compartment, and replace the batteries with proper orientation. Tactual cues on assistive devices and batteries are necessary to make this possible.

Table 1. Commercial Batteries

Type	General characteristics	Package	Capacity (AH)*	Cell V	WH/lb	WH/in^3	Operating temperature, °F (°C)	Shelf life (for 80% capacity)	Applications
	Primary Cells								
Zinc-air	Highest energy density on continuous discharge; excellent shelf life (unactivated); limited rate capability and shelf life when activated; flat discharge characteristic	Button Larger	1.15 8.5	1.4	140	19.0	32 - 122 (0 - 50)	5-10 years	For frequent use: Hearing aids, medical monitoring, pagers, communication equipment, data loggers, emergency lighting
Lithium-sulfur dioxide	High energy density; excellent high-rate and low-temperature performance; pressurized, hermetically sealed cell; excellent shelf life when activated; flat discharge characteristic	Cylindrical	0.7-19	3.0	125	7.1	-40 - 160 (-40 - 71)	5-10 years	High capacity, high-rate and extreme-temperature operation: Military and industrial applications
Lithium-manganese dioxide	High energy density; good rate capability and low-temperature performance; excellent shelf life; relatively flat discharge characteristic	Coin Cylindrical Bobbin	0.5 1.25 2.5	3.0	80 105 135	8.3 8.3 11.5	-4 - 140 (-20 - 60) -4 - 160 (-20 - 71)	5-10 years	General purpose applications requiring small high-capacity batteries: Watches, calculators, computers, memory backup, photoflash, camera motors
Alkaline	Popular, general-purpose premium battery; good low-temperature and high-rate performance; good shelf life; sloping discharge characteristic	Small Button Cylindrical	 20	1.5	25 60	2.4 5.2	-4 - 130 (-20 - 54)	3-4 years	General purpose high-drain applications: Lighting, cameras, tape recorders
Zinc-carbon	Popular, common, low-cost primary battery; moderate shelf life; sloping discharge characteristic	Cylindrical	40	1.5	34	2.3	23 - 113 (-5 - 45)	1-2 years	General purpose applications: Lighting, radios, novelties
Mercuric oxide	High capacity per unit volume; good shelf life; flat discharge characteristic	Button Cylindrical Larger	0.04-13	1.35 or 1.40	50 55	7.3 7.4	15 - 130 (-9 - 54)	3-5 years	For steady output voltage: Hearing aids, medical and photographic equipment, communication devices, pagers, detection and emergency equipment
Silver oxide	High capacity per unit weight; good shelf life; flat discharge characteristic	Button	0.18	1.5	60	8.2	-4 - 130 (-20 - 54)	2-3 years	Small high-capacity batteries: Hearing aids, watches, photographic equipment, and special electronics
	Secondary Cells								
Nickel-cadmium	Sealed: Maintenance free; good high-rate, low-temperature performance; excellent shelf life	Button Cylindrical	0.5 10	1.2	10.6 45	1.0 1.4	-40 - 113 (-40 - 45)	3-6 months	Portable tools, appliances, televisions, computers, memory backup, electronics
Lead-acid	Sealed: Maintenance free; low cost; good float capability	Cylindrical Flat-Plate	2.5-30 50	2.0	13.5 15	1.5 1.3	-40 - 140 (-40 - 60)	6-12 months	Portable tools, appliances, televisions, computers, electronics

* Button cells are rated at the 500-1000-hour rate at 70°F (21°C), cylindrical cells at the 50-100-hour rate. The cutoff voltage is approximately 80% of the nominal voltage. Typical capabilities may be higher.

Other issues that should be considered in the choice of batteries and battery compartments for assistive devices are safety, temperature range, reliability, energy density, and capacity for recharging. Safety issues of some batteries include leakage, toxicity, and risk of explosion if a short circuit occurs and venting fails. Many applications must warn of battery failure, because a warning device that is inoperative would give a false sense of security. It may be difficult for a sensory-impaired individual to detect the warning signs of a battery that is inoperative or about to leak or explode. Temperature range, reliability, size and weight are critical for devices that are in constant use, and frequent use may also indicate the use of rechargeable batteries, preferably with charging circuitry built into the assistive devices.

9.0 ADVANCED TECHNOLOGIES

Advanced technologies for portable power systems were discussed in Section 5.0. In general, zinc-carbon batteries are best for applications that do not require high current, extremes of temperature, or a high duty cycle. Zinc chloride batteries provide better performance with higher currents or greater duty cycle. Alkaline batteries are the next step up, for higher currents or duty cycles, and higher cost per battery. Military applications have used aluminum- or magnesium-based Leclanche cells for continuous drains, but these are being replaced by lithium batteries. For higher energy densities, zinc-mercuric oxide or zinc-silver oxide batteries may be used, but they cannot supply high currents, and their higher energy densities are accompanied by higher prices. Zinc-air cells provide even higher energy density, but they are unsuitable for intermittent use and have the limited temperature range of zinc-carbon batteries. Lithium batteries can achieve high energy density without the power limitations and limited intermittent-drain performance of other high-energy-density batteries. Their limitations, however, include cost, and, in some cases, safety considerations.

Secondary (rechargeable) batteries include lead-acid batteries, which are generally used when high capacity is needed, and nickel-cadmium batteries, which are used when a moderate capacity is appropriate. For higher performance at higher cost, for temperatures above 10°C (50°F), zinc-silver oxide batteries are available in a rechargeable configuration.

10.0 COST CONSIDERATIONS OF ADVANCED TECHNOLOGY

The best way to choose batteries for devices for persons with impairments, based on cost considerations, is to ensure commercial grade batteries are used, where possible, to meet fit, form and function. The choice of batteries for a given device depends upon the application and the user population.

11.0 COST BENEFITS TO PERSONS WITH SENSORY IMPAIRMENTS

Commercial or industrial batteries are almost always cheaper and more convenient, for persons with impairments, than special proprietary batteries, and their use should be strongly encouraged. Battery-intensive devices should use rechargeable batteries, if possible, and charging should be as effortless and automatic as the application permits.

12.0 PRESENT GOVERNMENT INVOLVEMENT

The Department of Education should continue to encourage commercial battery use; however, whenever possible, they should look at the possible use of military batteries, such as those used by NASA, which have higher current and longer life. The discussion presented above was based on commercial and industrial power supplies.

13.0 TECHNOLOGY TIMELINE

No battery development by the Department of Education is recommended.

14.0 PROPOSED ROAD MAP

The information provided in this scenario should be applied to the other scenarios in this report. This scenario should be disseminated to researchers and developers doing research for the Department of Education.

15.0 POTENTIAL PROGRAM SCHEDULE

The Department of Education should not undertake programs to develop better portable power options, because other departments of the U.S. Government, along with industry, are already engaged in these efforts and are better equipped for them. Instead, the Department of Education should encourage better use of battery technology as it is developed, including military batteries, to improve performance and/or reduce costs to the sensory-impaired population.